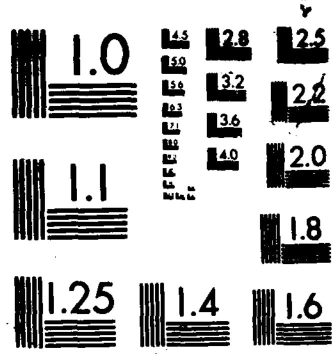
AD-A185 650 PROCEEDINGS OF THE DOD WORKLOAD ASSESSMENT WORKSHOP ON WORKLOAD ASSESSMEN..(U) NAVAL UNDERWATER SYSTEMS CENTER NEWPORT RI H M FIEDLER 15 SEP 87 NUSC-TD-6608 1/4

UNCLASSIFIED F/G 23/2 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

DTIC FILE COPY

(12)

NUSC Technical Document 6608
15 September 1987

Proceedings of the DoD Workload Assessment Workshop

H. M. Fiedler (Editor)

Naval Underwater Systems Center
Newport, Rhode Island / New London, Connecticut

3 2 81

87 11 ~~2 29~~

TABLE OF CONTENTS

AD-A185650

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) TD 6608	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Naval Underwater Systems Center 6b. OFFICE SYMBOL (If applicable) Code 2212	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) Newport Laboratory Newport, RI 02841-4057	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO / TASK NO. / WORK UNIT ACCESSION NO

11. TITLE (Include Security Classification)
PROCEEDINGS OF THE DoD WORKLOAD ASSESSMENT WORKSHOP

PERSONAL AUTHOR(S)
Fiedler, H.M. (Editor)

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
	FROM ______ TO ______	87-09-15	358

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Workload Assessment
05	05		Human Factors Engineering

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The Workload Subtechnical Advisory Group of the Department of Defense Human Factors Engineering Technical Advisory Group sponsored a workshop in "Workload Assessment Techniques and Tools," in September 1986 in Dayton, Ohio. This document is a compilation of the papers presented at the workshop.

DISTRIBUTION/AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☑ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL H.M. Fiedler	22b. TELEPHONE (Include Area Code) (401) 841-2648 22c. OFFICE SYMBOL Code 2212

DD FORM 1473, 84 MAR

83 APR edition may be used until exhausted
All other editions are obsolete

Workload Assessment Techniques and Tools

Sponsored by
The Workload Sub Technical Advisory Group
of
the Department of Defense Human Factors Engineering
Technical Advisory Group

27-28 September 1986
Dayton Convention Center
Dayton, Ohio

DTIC
COPY
INSPECTED
6

Chaired by
Heidi M. Fiedler
Naval Underwater Systems Center
(401)841-2648

DoD HFETAG Workload SubTAG Workshop Agenda
27-28 September 1986
Dayton, OH

SATURDAY 27 September:

8:00-9:00	Introduction to Workload*	T. Eggemeier
	Subjective Techniques	
9:00-12:00	SWAT*	G. Reid
12:00-1:30	Lunch	
1:30-2:30	Bi-Polar Technique*	S. Hart
	Task Analysis	
2:30-3:30	Task Analysis*	T. Ramirez
3:30-5:30	TASCO*	J. Armstrong

SUNDAY 28 September:

	Physiological Techniques	
9:00-10:00	NWTB*	G. Wilson
10:00-10:30	Break	
10:30-11:30	Eye Behavior and Evoked Response*	J. Stern
11:30-1:00	Lunch	
	Behavioral Tehcniques	
1:00-2:30	Behavioral Techniques*	W. Derrick/T. McCloy
2:30-3:00	Break	
	Other Tools	
3:00-4:30	Workload Index (WINDEX)*	B. North

*Tutorial Notes Included

Workshop Presenters

Mr. Jerry Armstrong
CSC-IFAST
Edwards AFB, CA 93523-5000
AV 527-6982

LT COL Thomas McCloy
DFBL
USAF Academy, CO 80840-5941
AV 259-3860, (303)472-3860

LT COL William Derrick
HQ USASFA/DFBL
USAF Academy
Colorado Springs, CO 80840-5941
AV 259-3175

Dr. John Stern
Department of Psychology
Washington University
St. Louis, MO 63104
(314)621-4211

Mr. Gary Reid
AAMRL/HEG
Wright-Patterson AFB, OH 45433-6573
AV 785-8749, (513)255-8749

Ms. Sandra G. Hart
ATTN: MS 239-3
NASA-Ames Research Center
Moffett Field, CA 94035
AV 359-6072

Dr. Glen Wilson
AAMRL/HEG
Wright-Patterson AFB, OH 45433-6573
AV 785-8748, (513)255-8748

Ms. Tammy Ramirez
JAYCOR
3164 Presidential Drive
Fairborn, OH 45324-2039
(513)429-4311

Dr. Robert North
Honeywell Inc.
3660 Technology Drive
Minneapolis, MN 55418
(612)782-7388

Mr. Thomas Eggemeier
Department of Psychology
University of Dayton
300 College Park Drive
Dayton, OH 45469
(513)229-2164

INTRODUCTION TO WORKLOAD

Presenters

F. T. Eggemeier: Workload Metrics for System Evaluation

F. T. Eggemeier and R. D. O'Donnell: A Conceptual Framework for Development of a Workload Assessment Methodology

Acquisition of Software

Copies of the analysis package for the SWAT, designed for use on an IBM PC, can be obtained by contacting:

Mr. Gary Reid
AAMRL/HEG
Wright-Patterson AFB, OH 45433-6573
AV 785-8749, (513)255-8749

Copies of the pencil and paper version or the computer version of the Bi-Polar Subjective Technique can be obtained by contacting:

Ms. Sandra G. Hart
ATTN: MS 239-3
NASA-Ames Research Center
Moffett Field, CA 94035
AV 359-6072

WORKSHOP OUTLINE

OVERVIEW OF WORKLOAD ASSESSMENT TECHNIQUES AND METRIC SELECTION CRITERIA

F. Thomas Eggemeier
University of Dayton

Background and Definitions

Major Classes of Empirical Assessment Techniques

- -Subjective Techniques
- -Performance-Based Techniques
  - -Primary Task Measures
  - -Secondary Task Methodology
- -Psychophysiological Techniques

Workload Metric Selection Criteria

- -Sensitivity
- -Diagnosticity
- -Intrusiveness
- -Implementation Requirements
- -Operator Acceptance

Metric Selection Guidlines

- -Type of Question to be Answered
- -Practical Constraints

REFERENCES

Chiles, W.D. Workload, task, and situational factors as modifiers of complex performance. In E.A. Alluisi & E.A. Fleishman (Eds.), Human Performance and Productivity. Hillsdale, New Jersey: Earlbaum Press, 1982.

Gartner, W.B., & Murphy, M.R. Pilot workload and fatigue: A critical survey of concepts and assessment techniques. Report No. NASA-TN-D-8365. Washington, D.C.: National Aeronautics and Space Administration, November, 1976.

Jahns, D.W. A concept of operator workload in manual vehicle operations. Report No. 14, Meckenheim, Germany: Forschungsinstitur fur Anthropotechnik, 1973.

Hart, S.G. Theory and measurement of human workload. In J. Zeidner (Ed.), Human Productivity Enhancement: Training and Human Factors in System Design., 1986, pp. 396-455.

Gopher, D., & Donchin, E. Workload: An examination of the concept. In K. Boff, L. Kaufman, & J. Thomas (Eds.), Handbook of Perception and Human Performance, Vol. II. New York: J. Wiley & Sons, Inc., 1986.

Moray, N. Subjective mental workload. Human Factors, 1982, 24, 25-40.

Moray, N. (Ed.), Mental Workload: Its Theory and Measurement. New York: Plenum Press, 1979.

O'Donnell, R.D. Contributions of psychophysiological techniques to aircraft design and other operational problems. AGARD Report No. AG-244, NATO Advisory Group for Aerospace Research and Development, 1979.

O'Donnell, R.D. & Eggemeier, F.T. Workload assessment methodology. In K. Boff, L. Kaufmann, & J. Thomas (Eds.), Handbook of Perception and Human Performance, Vol. II. New York: J. Wiley & Sons, 1986.

References (Continued)

Ogden, G.D., Levine, J.M., & Eisner, E.J. Measurement of workload by secondary tasks. Human Factors, 1979, 21, 529-548.

Rolfe, J.M. The secondary task as a measure of mental load. In W.T. Singleton, J.G. Fox, and D. Whitfield, (Eds.), Measurement of Man at Work. London: Taylor & Francis, 1971.

Wickens, C.D. Processing resources in attention. In R. Parasuraman & R. Davies (Eds.), Varities of Attention. New York: Academic Press, 1984.

Wierwille, W.W., & Williges, R.C. Survey and analysis of operator workload assessment techniques. Blacksburg, Virginia: Systemetrics, Inc., Report No. S-78-101, 1978.

Wierwille, W.W. & Williges, B.H. An annotated bibliography on operator mental workload assessment. Patuxant River, Maryland: Naval Air Test Center Technical Report No. SY-27R-80, March, 1980.

Workload Metrics for System Evaluation
by F. T. Eggemeier

PROCEEDINGS OF NATO DEFENSE RESEARCH GROUP PANEL VIII WORKSHOP,
"APPLICATIONS OF SYSTEM ERGONOMICS TO WEAPON SYSTEM DEVELOPMENT"
SHRIVENHAM, ENGLAND, 1984, pp. C-5 - C-20.

WORKLOAD METRICS FOR SYSTEM EVALUATION

F. Thomas Eggemeier

Wright State University

and

Systems Research Laboratories, Inc.
Dayton, Ohio

INTRODUCTION

A major function of human factors engineering throughout the system development process is to ensure that system demands do not exceed the information processing capabilities of the human operator. Processing overload is a central factor leading to breakdowns in operator performance and to the compromises in system safety and effectiveness that can result from such decrements. Mental workload is the term which refers to that portion of an operator's limited processing capacity which is actually required to perform a particular task or system function. The principal objective of workload assessment is to specify the amount of expended processing capacity so that existing or potential overloads can be identified and decrements in operator performance avoided.

The use of advanced display and control technologies in modern weapons systems has been accompanied in many instances by substantial increases in the monitoring, supervisory, and decision-making demands imposed on the operator. These heavy demands have markedly increased the likelihood of approaching or actually exceeding operator processing capacity limits. As a consequence, assessment of the mental workload imposed by alternative design options has become particularly critical throughout the weapon system design process.

Because of its critical role in the system development process, workload assessment has been the subject of considerable research over the past 10 years (e.g., Moray, 1979). One product of these research efforts has been the development and application of a large number of individual workload assessment techniques. A recent comprehensive review (Wierwille and Williges, 1978) of the workload assessment literature, for example, identified 28 different techniques that had been used to derive measures of load. A substantial number of these empirical assessment techniques can be classified as belonging to one of three categories of workload measures: (1) subjective opinion procedures, (2) performance-based techniques, and (3) physiological techniques.

Subjective techniques (e.g., Gartner and Murphy, 1976; Williges and Wierwille, 1979; Moray, 1982) require that the operator judge and report the degree of workload experienced during performance of a particular task or system function. Rating scales are the most frequently used type of subjective measurement technique.

Performance-based techniques (e.g., Gartner and Murphy, 1976; Williges and Wierwille, 1979) use some measure of operator behavior or activity as the basis of a workload index. A number of individual assessment techniques can be categorized as performance-based measures. So-called primary task techniques (e.g., Rolfe, 1976; Gartner and Murphy, 1976; Williges and Wierwille, 1979) examine some aspect of the operator's capability to perform the task or system function of interest in order to provide an estimate of load. Deviations from glideslope by a pilot on final approach would constitute one such primary task measure. A second type of performance-based measure which has been frequently used to assess workload is secondary task methodology (e.g., Knowles, 1963; Rolfe, 1971; Ogden, Levine, and Eisner, 1979; Williges and Wierwille, 1979). This approach derives an estimate of workload from the operator's capability to perform a secondary task concurrently with the primary task of interest.

Physiological techniques (e.g., O'Donnell, 1979; Wierwille, 1979) measure some aspect of the operator's physiological response to task or system demand, and provide a measure of load based on these responses. A wide variety of physiological measures (e.g., heart rate variability, pupil diameter, event-related brain potentials) have been used in order to assess workload.

Since a variety of workload assessment procedures are available, an important decision faced by a system designer involves choice of the technique that best meets design requirements. The system development process typically involves a series of stages which range from conceptual development through operational test and evaluation of the system. These stages can be characterized by variations in both the specific questions addressed by workload measurement, and in the practical constraints that must be satisfied by assessment techniques. These questions and constraints suggest a number of criteria that should be considered in choosing a workload measure for application during system development. The purposes of this paper are to outline a set of such criteria, briefly review the current status of the three classes of empirical techniques as they relate to the proposed criteria, and suggest some applications for each class of technique during system development. Some recent work with a subjective assessment procedure which has the potential for application throughout the system development process is also discussed.

WORKLOAD METRIC SELECTION CRITERIA

A number of criteria for evaluation of workload metrics have been proposed in the recent literature (e.g., Gartner and Murphy, 1976; Rolfe, 1976; Ogden et al., 1979; Williges and Wierwille, 1979; Wickens, 1981; Shingledecker, 1983). Several of the proposed criteria are particularly relevant for choice of a metric during system design. These criteria include: (1) sensitivity, (2) diagnosticity, (3) intrusiveness, (4) implementationn requirements, and (5) operator acceptance.

Sensitivity

Sensitivity refers to the capability of a measure to distinguish different levels of load imposed by a task or design option. The degree of sensitivity required in an assessment technique is directly related to the nature of the question to be answered by the workload measure. There are a wide variety of specific design questions (e.g., adequacy of control/display design, allocation functions between operators) that can be addressed by workload assessment during system development. Regardless of the specific aspect of the design that is addressed, however, the two basic objectives of workload assessment are to determine: (1) if an overload that would lead to degraded operator performance actually exists, or (2) if the potential for such an overload exists. Questions involving the first objective can be addressed through primary task performance measures, since they are generally assumed to differentiate overload from nonoverload situations (e.g., Knowles, 1963; Gartner and Murphy, 1976; Williges and Wierwille, 1979). In other applications, however, a designer might wish to evaluate the potential for overload among several design options that yield adequate operator performance. This objective is relevant when it is anticipated that other factors during system operation (e.g., environmental stressors, equipment failures) might contribute additional load that would be sufficient to cause degraded operator performance. In this instance, even though none of the design options themselves overload the operator, it is desirable to identify the option that imposes the lowest load and affords the greatest reserve capacity for dealing with other sources of demand. This type of evaluation would require a workload measure that was more sensitive to variations in load than primary task measures, and would suggest the use of other procedures (e.g., subjective, physiological, secondary task) that are designed to discriminate levels of workload in nonoverload situations. Current evidence indicates, for example, that both secondary task measures (e.g., Schifflet, Linton, and Spicuzza, 1982) and subjective ratings of load (e.g., Eggemeier, Crabtree, and LaPointe, 1983) can discriminate differences in task demand that are not reflected in primary task measures of operator performance. The sensitivity criterion is, therefore, an essential consideration in choice of a workload measure, since the degree of sensitivity bears directly on the type of question that can be addressed by a technique.

Diagnosticity

Diagnosticity (Wickens, 1981; Wickens and Derrick, 1981; Shingledecker, 1983) is a second important consideration in choice of a system evaluation metric. This criterion is based on the multiple resources theory (e.g., Navon and Gopher, 1979; Sanders, 1979; Wickens, 1981) explanation of limitations within the human processing system. Essentially, this theory holds that the processing capacity expended in task performance is not unitary, but is drawn from multiple sources or pools, each with its own resources that cannot be exchanged with other pools. One version of multiple resources theory (Wickens, 1981) maintains that perceptual and central processing stages within the human system draw on one resource pool, while the response or motor output stage draws from a separate resource pool. Under this position, it is possible to overload or fully expend the resources associated with one source, while not depleting the processing resources of another source. For example, the requirement to monitor a display which places heavy demands on short-term memory might overload perceptual/central processing resources, while making minimal demands on motor output resources. Other system requirements such as a final approach in an aircraft would have a different demand composition, and might require greater expenditures of motor output resources. Diagnosticity refers to the capability of a technique to discriminate these differences in the load imposed on specific operator resources.

It has been proposed (Wickens, 1981; Wickens and Derrick, 1981) that workload measures vary in their degree of diagnosticity. There are data which indicate, for example, that some physiological measures such as pupil diameter (e.g., Beatty and Kahneman, 1966; Jiang and Beatty, 1981) and some subjective rating scales (e.g., Reid, Shingledecker, and Eggemeier, 1981a; Eggemeier, Crabtree, Zingg, Reid, and Shingledecker, 1982; Notestine, 1983; Wierwille and Casali, 1983a) are sensitive to perceptual, central processing, and response load manipulations. The event-related brain potential (Isreal, Chesney, Wickens, and Donchin, 1980; Isreal,, Wickens, Chesney, Donchin, 1980) and some secondary tasks (e.g., North, 1977; Wickens and Kessel, 1980; Shingledecker, Acton, and Crabtree, 1983), however, show differential sensitivity to manipulations of perceptual/central processing and motor output demands. These data imply that subjective rating scales and some physiological measures are not particularly diagnostic, and can prove sensitive to variations in resource expenditure anywhere within the human processing system. However, other physiological metrics and various secondary tasks appear to be more diagnostic of specific types of resource or capacity expenditure.

Such differences in diagnosticity suggest that the different types of measures can play complementary roles during system development. Less diagnostic measures could serve as screening devices to initially determine if high levels of loading exist during performance of a task or system function, while more diagnostic procedures could be subsequently used to pinpoint the particular source (e.g., perceptual versus motor output) of any such overloads. Choice of an assessment technique on the basis of the diagnosticity criterion would, therefore, be dependent on the objective to be met by the measure of workload.

Intrusiveness

While the criteria of sensitivity and diagnosticity relate to the nature of the question that is to be addressed by a workload measure, there are a number of additional criteria that are suggested by practical constraints imposed on the use of metrics during the system development process. The characteristic of intrusiveness (e.g., Gartner and Murphy, 1976; Williges and Wierwille, 1979; Shingledecker, 1983) is one such criterion, and refers to the tendency for some metrics to cause degradations in ongoing primary task performance.

Intrusiveness in an assessment procedure is undersirable on both practical and theoretical grounds. From a practical perspective, it is clear that any technique that causes decrements in operator performance can potentially compromise the safety of system operation. Such compromises are obviously unacceptable, particularly during the later stages of system development when operational test and evaluations of prototype or initial production models are conducted. From a theoretical point of view, intrusiveness can cause problems in the interpretation of data resulting from application of an assessment technique. These interpretation problems stem from the assumption that measurement procedures provide a pure index of the load imposed by the primary task. If primary task performance is degraded by the introduction of the assessment technique, an unbiased measure of primary task workload is not possible. Although intrusiveness presents potential difficulties for all metrics, the interpretation problem can be particularly acute with secondary task measures (Rolfe, 1971; Ogden et al., 1979) that are intended to provide a measure of the reserve capacity afforded by the primary task.

Despite its importance, the comparative data base on the degree of intrusion associated with the various types of metrics is not extensive. Some significant steps toward establishing a systematic data base have been undertaken recently (e.g., Casali and Wierwille, 1982, 1983; Rahimi and Wierwille, 1982; Shingledecker, Crabtree, and Acton, 1982; Acton, Crabtree, and Shingledecker, 1983; Wierwille and Casali, 1983b; Wierwille and Conner, 1983), but such direct comparison data are not yet complete. However, some statements regarding the potential for intrusiveness can be made on the basis of data generated by individual applications of the various techniques.

First, it is clear that intrusiveness has represented a major problem in many applications of secondary task methodology (e.g., Rolfe, 1971; Gartner and Murphy, 1976; Ogden et al., 1979; Williges and Wierwille, 1979). The problem has led to the development of techniques such as cross-adaptive (e.g., Kelly and Wargo, 1967; Jex and Clement, 1979) and embedded (Shingledecker, Crabtree, Simons, Courtright, and O'Donnell, 1980; Shingledecker, 1980; Crabtree and Spicuzza, 1981; Shingledecker and Crabtree, 1982) secondary tasks that are designed to minimize or control the levels of intrusion. Cross-adaptive procedures permit variations in secondary task difficulty as a function of primary task performance. When primary task performance falls below a specified criterion, secondary task difficulty is reduced in order to control the level of intrusion. This type of procedure has been successfully employed in a number of laboratory and simulation studies (Kelly and Wargo, 1967; Jex and Clement, 1979) that have utilized primary continuous tracking tasks. Applications of the procedure to discrete tasks in more complex environments have not been accomplished, and could present difficulties due to problems in obtaining primary task measures that would permit adaptation of the secondary task. The embedded secondary task approach, on the other hand, was developed for application to high fidelity simulation or operational environments. This procedure uses an element already embedded in normal system operation procedures as the secondary task. The elements chosen as secondary tasks (e.g., radio communications) are those that are normally assigned lower priority than the primary task (e.g., flight control), thereby minimizing the potential for primary task intrusion.

Second, it appears that the intrusion associated with most other classes of assessment techniques tends to be minimal. Subjective assessment techniques typically present no significant intrusion problem, sir e rating scales and other report procedures are usually completed subsequent to primary task performance. Primary task measures are, by definition, nonintrusive, because their application involves no additional operator performance or reports. Physiological procedures also appear to minimize the potential for intrusion, although there are data (Rahimi and Wierwille, 1982) which indicate that these techniques can be associated with some intrusion.

The degree of intrusiveness that can be tolerated in an assessment technique will vary as a function of the context in which the measure is taken. Some degree of intrusion in a simulator or in a crewstation mockup could be less serious, for example, than equivalent levels of primary task decrement during actual system operation. Choice of an assessment procedure on the basis of intrusiveness would, therefore, be determined in part by constraints dictated by the measurement situation.

Implementation Requirements

The implementation requirements associated with a particular measurement technique constitute a second criterion that is heavily influenced by the practical constraints imposed by the system development process. Implementation requirements are factors that are related to the ease with which a technique can be applied at different stages of system development and evaluation. Examples of such factors include: (1) the instrumentation and software that is required to record and analyze the measures associated with a technique; (2) any operator training that is necessary for the technique to be properly applied; and (3) system simulation facilities or actual equipment that are required for application of the technique.

Different classes of assessment procedures can vary considerably in their instrumentation requirements, as can individual techniques within the same category. For instance, subjective opinion measures usually make use of paper and pencil for data recording, while much more stringent implementation requirements are typically associated with physiological and some performance-based procedures. Requirements also vary within categories themselves. Cross-adaptive secondary techniques require more extensive instrumentation than other secondary task procedures (e.g., interval production, Shingledecker et al., 1983) which require only a means of recording an operator's response. Therefore, when minimal instrumentation is a primary constraint, the use of subjective measures or certain secondary task procedures such as the interval production task is suggested.

Operator training requirements also vary with techniques and can be necessary with both secondary task and subjective assessment procedures. Applications of secondary task methodology, for instance, usually require some operator training in order to stabilize baseline performance on the secondary task before it is performed concurrently with the primary task. Some subjective procedures (e.g., Reid et al., 1981a) also include the provision for familiarization with the rating scales prior to their use. Training requirements associated with the use of both primary task and physiological measures would be virtually nonexistent in most cases.

Techniques can also differ in the types of simulation facilities and operational equipment that are necessary for their application. Such facility requirements can be particularly restrictive during the early conceptual stage of system development, when system design information is very general, and simulation and mockup facilities are typically not available. Since both performance-based and physiological techniques require such facilities, their application has been usually restricted to later stages (e.g., validation, engineering development) of the design process when the appropriate devices are present. This constraint on early use of physiological and performance-based procedures is one factor that has led to the development and application of analytical time-line techniques (e.g., Zipoy, Premselaar, Gargett, Belyea, and Hall, 1970; Parks, 1979; Geer, 1981) and several simulation models (e.g., Linton, Jahns, and Chatelier, 1977; Lane, Strieb, and Wherry, 1977; Lane, Strieb, Glenn, and Wherry, 1981; Chubb, 1981) that are capable of addressing workload assessment issues during earlier stages of design. Traditional applications of subjective metrics also require the availability of mockups, simulators, or operational equipment. However, a recent application (Quinn, Jauer, and Summers, 1982) demonstrated the projective use of a subjective metric by requiring experienced pilots to rate the expected load associated with several proposed cockpit enhancements. The projective ratings were based on detailed descriptions of mission profiles and control/display options, and were intended to provide workload estimates that could be combined with other factors (e.g., cost) to initially screen design options for further evaluation. Although the results must be validated, the Quinn et al. study provides a methodology with the potential to permit application of subjective procedures during the earlier stages of development when performance-based and physiological techniques are not practicable.

Taken together, implementation requirements can therefore impose important constraints on the use of the various classes of assessment techniques during the development process. Instrumentation and facility requirements are typically more stringent with performance-based and physiological techniques than with subjective procedures, suggesting the use of the latter for certain situations.

Operator Acceptance

The characteristic of operator acceptance is important to ensure that an assessment technique will yield data that are representative of the load imposed by the task or system function in question. Assessment procedures which are perceived by operators as bothersome or artificial incur the risk of being ignored, performed at substandard levels, or being associated with significant levels of primary task intrusion. Any of these factors can lead to compromises in the effectiveness of a technique.

In spite of the potential importance of operator acceptance, there are little or no formal comparative data which are available to address operator reaction to the major classes of techniques. Although some investigators (e.g., Hallsten and Borg, 1975) have commented on operator acceptance of a number of procedures, the data are not sufficient to address the issue in a comprehensive manner. Informal data and knowledge of the procedures involved in application of the techniques can, however, be used to provide some estimates of acceptance. Informal evidence, for example, suggests that subjective procedures usually enjoy a high degree of user acceptance, quite possibly because of the high face validity associated with many current rating scales (e.g., Cooper and Harper, 1969; Reid et al., 1981a). Operator acceptance should also be quite good for primary task measures, since they do not typically involve any additional operator response or effort. Physiological techniques would have some potential for low acceptance if the recording instruments used are considered bothersome by the operator, but this does not appear to have been a significant problem with most techniques. Secondary task methods could also be considered distracting by the operator if the requirement to perform the secondary task interferes with primary task performance. The embedded secondary task technique (Shingledecker et al., 1980) which utilizes a secondary task that is normally performed in the operational environment should, however, minimize this risk.

APPLICATION GUIDELINES

It is obvious from the foregoing discussion that no single assessment technique is capable of meeting all of the criteria outlined above. The various categories of techniques are characterized by the capability to satisfy some criteria, but not others. Criteria vary in their importance as a function of the different stages of design, and consequently, techniques vary in their applicability. It is therefore clear that assessment of workload across the various phases of the design process will require the complementary use of multiple metrics, since no single metric is capable of providing all of the required information.

The capability of individual assessment procedures to meet the various criteria can provide some guidance regarding their use for specific purposes at different stages of design. Table 1 summarizes the current status of the procedures with respect to the proposed criteria, and can be used as a basis to suggest particular applications for each class of technique.

An investigator requiring a nonintrusive general measure of load in an operational environment with restricted data recording capabilities should, for example, consider the application of subjective metrics. On the other hand, primary task measures might be considered for application in a high fidelity simulator with performance measurement capability when the objective was to evaluate the adequacy of operator performance with a particular design option. The use of secondary task methodology or an appropriate physiological technique in a system simulator would be suggested if the intent was to isolate the source of

TABLE 1. SUMMARY OF WORKLOAD ASSESSMENT TECHNIQUE CAPABILITIES

	Sensitivity	Diagnosticity	Intrusiveness	Implementation Requirements	Operator Acceptance
PRIMARY TASK MEASURES	Discriminate overload from nonoverload situations. Used to determine if operator performance will be acceptable with a particular design option.	Not considered diagnostic. Represents a global measure of workload that is sensitive to overloads anywhere within the operator's processing system.	Nonintrusive since no additional operator performance or report required.	Instrumentation for data collection can restrict use in operational environments. Use requires mockups, simulators, or operational equipment. Imposes limits on use during early system development. No operator training required.	No systematic data. No reason to expect negative operator opinion.
SECONDARY TASK METHODS	Capable of discriminating levels of capacity expenditure in nonoverload situations. Used to assess reserve capacity afforded by a primary task. Can be used to assess the potential for overload among design options.	Capable of discriminating some differences in resource expenditure (e.g., central processing versus motor). Diagnosticity suggests complementary use with more generally sensitive measures, with the latter initially identifying overloads and secondary tasks being used subsequently to pinpoint the locus of overload.	Primary task intrusion has represented a problem in many applications, particularly in the laboratory. Data are not extensive in operational environments. Several techniques (e.g., embedded secondary task, adaptive procedures) have been designed to control intrusion. Potential for intrusion could limit use in operational environments.	Instrumentation for data collection can restrict use in operational environments, but some tasks have been instrumented for in-flight use. Use requires mockups, simulators, or operational equipment. Imposes limits on use during early system development. Some operator training usually required to stabilize secondary task performance.	No systematic data. Requirement to perform secondary task could distract operator. Technique such as embedded secondary task should minimize any acceptance problems.
PHYSIOLOGICAL TECHNIQUES	Capable of discriminating levels of capacity expenditure in nonoverload situations. Can be used to assess the relative potential for overload among design options.	Some techniques (e.g., event-related brain potential) appear diagnostic of some resources, while other measures (e.g., pupil diameter) appear more generally sensitive. Choice of technique dependent on purpose of measurement (screening for any overload versus identifying locus of overload).	Intrusion does not appear to represent a major problem, although there are data to indicate that some interference can occur.	Instrumentation for data collection can restrict use in operational environments. Use requires mockups, simulators, or operational equipment. Imposes limits on use during early system development. No operator training required.	No systematic data. Instrumentation and recording equipment could represent potential problem, but no significant problems reported in literature.
SUBJECTIVE TECHNIQUES	Capable of discriminating levels of capacity expenditure in nonoverload situations. Can be used to assess the relative potential for overload among design options.	Not considered diagnostic. Available evidence indicates that rating scales represent a global measure of load. Lack of diagnosticity suggests use as a general screening device to determine if overload exists anywhere within task performance.	Intrusion does not appear to represent a significant problem. Most applications require rating scale completion subsequent to task performance and, therefore, present no intrusion problem.	Instrumentation required is usually minimal, permitting use in a number of environments. Traditional applications require mockups, simulators, or operational equipment. Imposes limits on use during early system development. Recent projective use provides potential for application during early stages. Some familiarization with procedures can be required.	No systematic data. Informal evidence suggests that several rating scales enjoy a high degree of operator acceptance.

an overload that had been previously identified through use of a subjective or primary task metric. The potential applications for each class of metric are clearly much more extensive than those suggested by these hypothetical situations. The examples do, however, illustrate how the proposed criteria can be applied to identify the class(es) of techniques that might be most appropriate for a particular application.

In many instances, use of the proposed criteria will result in simultaneous application of more than one technique. Applications of secondary task methodology, for example, require the measurement of primary task performance in order to evaluate the degree of any intrusion that might have occurred. In other instances, the objectives of an evaluation might also suggest the concurrent use of more than one metric. For example, a comprehensive evaluation of two display options might include the use of both primary task and subjective measures. The primary task measure would permit assessment of any differences in the adequacy of task performance that could be expected with the options, while the subjective technique would provide the potential to identify any workload differences between the options that were not reflected in the less sensitive performance measure.

The preceding review and discussion of metrics has been primarily concerned with classes of workload assessment techniques in general. It is clear from the foregoing discussion, for example, that the general category of subjective metrics holds a great deal of potential for use during system design. Once a class of technique has been identified as appropriate, however, an individual procedure or measure from within the category must be chosen for actual application. Individual procedures themselves can also vary along a number of dimensions that can impact their suitability for use. The purpose of the following discussion is to briefly review some recent work with an individual subjective assessment technique that appears to be particularly well suited for a number of applications throughout the system development process.

SUBJECTIVE WORKLOAD ASSESSMENT

Subjective workload measurement procedures satisfy a number of the criteria outlined above and, as a consequence, have been very frequently employed as workload assessment techniques (e.g., Williges and Wierwille, 1979). Despite their advantages, there are several problems which have been traditionally associated with use of subjective workload metrics. First, in many applications, individual rating scales have been developed for a specific investigation and have not been validated for generalized use. Second, there is little evidence in the literature of workload rating scales that have been rigorously developed on the basis of psychometric procedures (e.g., Williges and Wierwille, 1979). As a consequence, most available rating scales have unknown metric properties, and must be assumed to provide only ordinal level measurement.

In order to provide a workload rating scale with known metric properties and with the potential for generalized applicability, a procedure termed the Subjective Workload Assessment Technique (SWAT) has been developed (Reid et al., 1981a; Reid, Shingledecker, Nygren, and Eggemeier, 1981b; Reid, Eggemeier, and Shingledecker, 1982). In SWAT, it is assumed that there are three major contributors to subjective mental load: (1) time load, (2) mental effort load, and (3) psychological stress load. Time load refers to the percentage of time that an operator is busy, and reflects such factors as overlap and interruption among tasks. Mental effort load, on the other hand, refers to the degree of attention or concentration required during task performance. The final dimension, psychological stress load, reflects any additional factors that cause operator anxiety or confusion and, therefore, contribute to subjective mental load. In SWAT, each of the three dimensions is represented by an individual three-point rating scale with verbal descriptors that define the levels on each dimension.

SWAT is based on application of conjoint measurement and scaling (e.g., Nygren, 1982). Conjoint measurement and scaling permit ratings on the three dimensions to be combined into one overall scale of workload with interval measurement properties. In order to identify the rule which is appropriate for combining the three dimensions into the overall interval scale, a scale development phase is completed. During this phase, subjects rank-order the subjective load associated with the 27 possible combinations that result from the three levels of time, mental effort, and psychological stress load. This rank-ordering information is subjected to a series of axiom tests to identify the rule for combining the three dimensions. When the rule has been established, conjoint scaling is applied to derive the overall scale of workload. Subsequent to the scale development phase, subjects participate in an event scoring phase. During event scoring, subjects perform the task or mission segment of interest and rate the time, mental effort, and stress load associated with performance. The ratings on the individual dimensions are then converted to one of the 27 points on the interval scale that was derived during scale development. More extensive discussions of the scale development and event scoring procedures can be found in Reid et al. (1981a,b), and Reid, Eggemeier, and Nygren (1982).

One aspect of the work conducted during the development of SWAT has centered on establishing its capability to reflect workload differences in a number of different types of tasks in several environments that are representative of those found during system development. SWAT has been successfully applied in a number of laboratory or part-task simulation environments (e.g., Reid et al., 1981a; Eggemeier et al., 1982, 1983; Notestine, 1983); in several full mission simulators (e.g., Reid, Eggemeier, and Shingledecker, 1984; Skelly, Reid, and Wilson, 1983); and under conditions that are similar to the early stages of system development when workload estimates must be based on detailed mission scenarios and descriptions of system equipment capabilities (Quinn et al., 1982).

Figure 1 illustrates the results of two applications of SWAT in laboratory/part-task simulation environments. Panel A (Reid et al., 1981a) shows the results of an experiment which employed several levels of a simulated flight control (critical tracking, Jex and Clement, 1979) task and a secondary simulated aircrew radio communications task (Shingledecker et al., 1980). Significant differences in SWAT ratings were obtained in the communication task alone condition versus the more difficult dual task condition. SWAT ratings also successfully discriminated levels of difficulty in both the simulated flight control and radio communications tasks. Panel B (Eggemeier et al., 1983) illustrates the effects on SWAT ratings of variations in the rate of stimulus presentation in a sequential short-term memory task. Subjects in the experiment were required to monitor a visual display and update the status of four categories of information that changed at several rates. The memory task was intended to be representative of the demands placed on air traffic controllers while monitoring flight control displays. SWAT ratings successfully discriminated levels of difficulty in the memory task, even though a primary task measure of performance errors showed no significant differences between conditions.

Several recent experiments also support the applicability of SWAT to full mission simulation environments. SWAT ratings have proven sensitive to expected workload variations in high fidelity flight simulation evaluations of advanced control/display options in both fighter (Reid et al., 1984) and bomber (Skelly

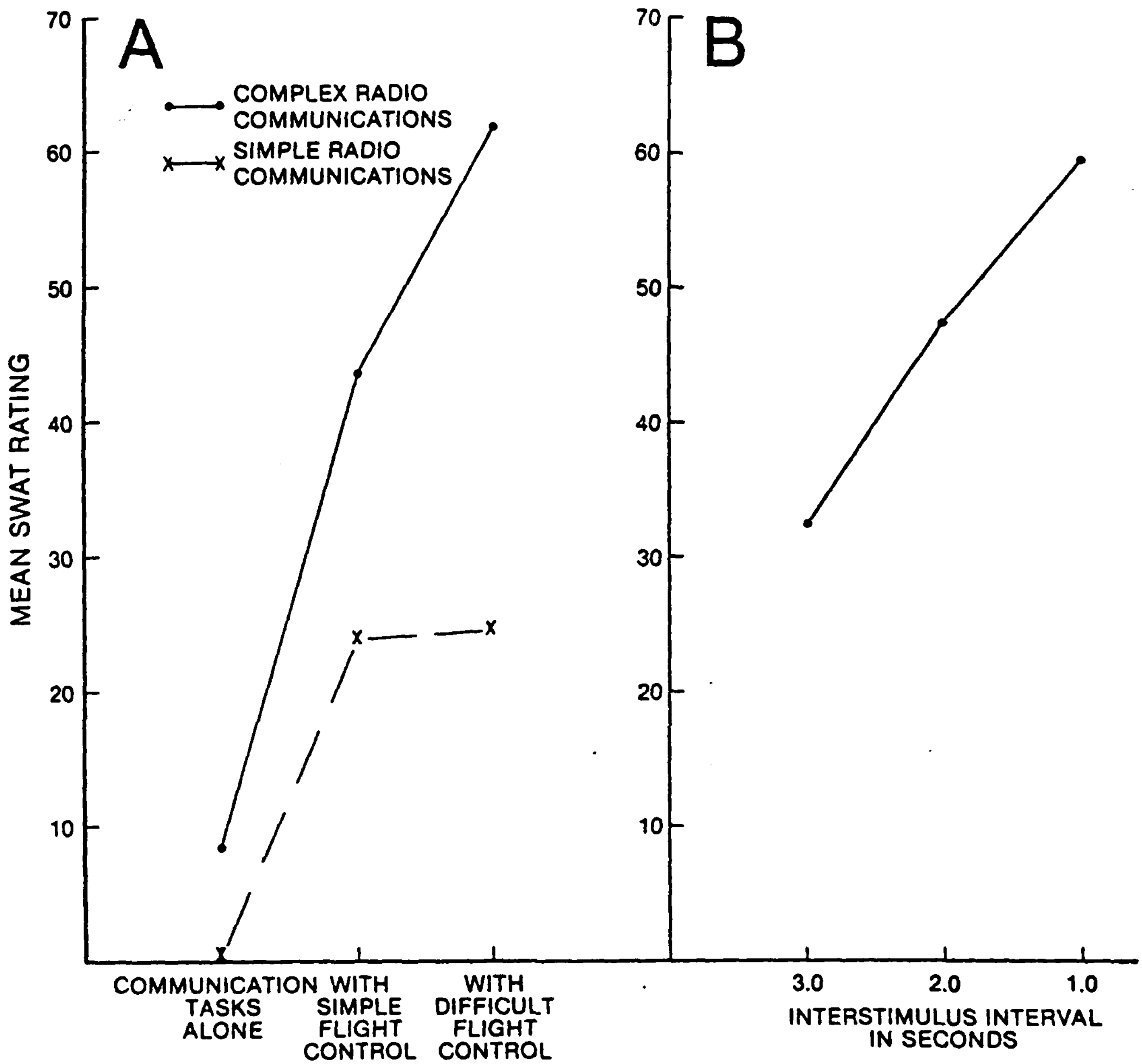


Figure 1. Mean SWAT Ratings as a Function of Task Difficulty in Two Experiments. [Panel A illustrates the effects of simple and complex radio communications on SWAT ratings in both single and dual task conditions (Figure drawn from the data of Reid et al., 1981a). Panel B shows the effect of stimulus presentation rate manipulations in a sequential short-term memory task (Figure adapted from Eggemeier et al., 1983).]

et al., 1984) aircraft. Reid et al., for example, obtained significant differences in pilot SWAT ratings as a function of variations in the number of opponents during a fighter mission. SWAT ratings in the Skelly et al. study also showed differences that were logically defensible and consistent with expectations. Pilot ratings, for instance, were generally higher than copilot ratings, except for a number of segments in which the copilot was flying the aircraft. Segments of simulated mission which included various types of threats to the aircraft were rated higher than baseline segments that did not include such threats. In both applications, pilot acceptance of the rating procedure was very high, and in both instances, SWAT ratings were taken with minimal intrusion by having the pilot verbally report ratings after completion of a mission segment to an experimenter stationed outside the cockpit.

The Quinn et al. (1982) experiment that was briefly discussed earlier also utilized the SWAT methodology in a novel application of the technique. The purpose of the Quinn et al. study was to evaluate a variety of methods for enhancing fighter aircraft systems, including advanced display, control, and navigational concepts. A methodology was devised to comparatively evaluate the enhancements along a number of dimensions prior to prototype development, and SWAT was included to quantify predicted effects on pilot workload. A number of experienced fighter pilots were provided with a mission scenario and detailed descriptions of an advanced baseline version of the aircraft and several enhancements. On the basis of the information, the pilots provided mission SWAT ratings for the various versions of the baseline system that included several combinations of enhancements. The interval level data that are obtained from the SWAT procedure permitted use of the resulting workload ratings in a multiattribute utility analysis with other factors (cost, system performance) to permit selection of several options for further research. Although it is clear that the results of the projective SWAT ratings must be validated, the methodology employed by Quinn et al. is significant in that it demonstrates the feasibility of obtaining SWAT ratings on the basis of detailed mission and equipment information. Use of the technique in this manner includes obvious time and cost advantages and, as noted previously, demonstrates the potential for application of SWAT during the earlier stages of system design.

Taken together, the results of current work with the SWAT technique clearly support its sensitivity to a variety of tasks that are relevant to system operation. The available evidence also indicates that SWAT has a very high potential for applicability across several stages of design. These data, coupled with the advantages of the interval level measurement afforded by the technique, strongly support the utility of the SWAT metric for evaluation of workload during the system development process.

SUMMARY AND CONCLUSIONS

Application of the proposed criteria to the major categories of workload assessment techniques indicates that a battery of performance-based, subjective, and physiological metrics will be required to meet the varied needs for workload measurement that arise during the system development process. In many instances, the capabilities of one technique supplement those of another procedure, suggesting the complementary use of the various metrics at different stages of design. Among the classes of assessment procedures reviewed above, subjective techniques appear to have the greatest potential for application across the various phases of the design process, and the SWAT technique is one such procedure that has demonstrated high levels of sensitivity and applicability.

Although current information is sufficient to suggest some applications for the various categories of techniques, more extensive data are needed to refine proceuures for choice of a metric for particular applications. For example, more complete comparative data on the relative sensitivity and intrusiveness that can be expected from individual techniques from within particular categories (e.g., secondary task) of procedures represent a need in this area. As was noted previously, several such efforts have been recently undertaken (e.g., Shingledecker et al., 1983; Wierwille and Casali, 1983b), and the results should provide a more refined basis for choice of metric for particular applications. An additional area requiring further experimentation deals with the extension of current classes of metrics to the earlier and later stages of the design process. Implementation requirements have somewhat limited the applicability of secondary task and physiological metrics in the early and latter stages of system design, and more work is required to evaluate the application of these techniques beyond the laboratory and simulation environments. Some of this type of work (e.g., Schifflet et al., 1982) has been conducted, but additional efforts are required. Further evaluation and extension of the Quinn et al. (1982) procedure for application of subjective metrics during the early stages of design should also be pursued in order to supplement available analytic and modeling procedures that provide the current capability for workload assessment during this phase of the development process.

Release

This paper has been approved for public release, distribution unlimited.

Acknowledgement

This work was partially supported by the USAF Aerospace Medical Research Laboratory under Contract No. F33615-82-C-0511. Mr. Mark S. Crabtree provided very helpful comments on an earlier version of this paper.

REFERENCES

Acton, W. H., Crabtree, M. S., and Shingledecker, C. A., 1983, Development of a standardized workload metric evaluation methodology, Proceedings of the 1983 IEEE National Aerospace and Electronics Conference, 1086-1089.

Beatty, J. and Kahneman, D., 1966, Pupillary changes in two memory tasks, Psychonomic Science, 55:371-372.

Casali, J. G. and Wierwille, W. W., 1982, A sensitivity/intrusion comparison of mental workload estimation techniques using a flight task emphasizing perceptual piloting activities, Proceedings of the 1982 IEEE International Conference on Cybernetics and Society, 598-602.

Casali, J. G. and Wierwille, W. W., 1983, Communications-imposed pilot workload: A comparison of sixteen estimation techniques, Proceedings of Second Ohio State University Symposium on Aviation, 223-235.

Chubb, G. P., 1981, SAINT, A digital simulation language for the study of manned systems: in: "Manned Systems Design Methods, Equipment, and Applications," J. Morssel and K. F. Kraiss, eds., Plenum Press, New York.

Cooper, G. E. and Harper, R. P., Jr., 1969, The use of pilot rating in the evaluation of aircraft handling qualities, Report No. NASA TN-D-5513, National Aeronautics and Space Administration, Ames Research Center, Moffett Field, California.

Crabtree, M. S. and Spicuzza, R. J., 1981, Evaluation of embedded radio communications activities as secondary tasks for objective assessment of aircrew workload in simulators, trainers, and actual systems. Proceedings of the 1981 IEEE National Aerospace and Electronics Conference, 1349-1352.

Eggemeier, F. T., Crabtree, M. S., Zingg, J. J., Reid, G. B., and Shingledecker, C. A., 1982, Subjective workload assessment in a memory update task, Proceedings of the 1982 Human Factors Society Annual Meeting, 643-674.

Eggemeier, F. T., Crabtree, M. S., and LaPointe, P. A., 1983, The effect of delayed report on subjective workload ratings, Proceedings of the 1983 Human Factors Society Annual Meeting, 139-143.

Gartner, W. P. and Murphy, M. R., 1976, Pilot workload and fatigue: A critical survey of concepts and assessment techniques, Report No. NASA-TN-D-8365, National Aeronautics and Space Administration, Washington, D.C.

Geer, C. W., 1981, Human engineering procedures guide, Technical Report No. AFAMRL-TR-81-35, USAF Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio.

Hallsten, L. and Borg, G., 1975, Six rating scales for perceived difficulty, Report #58 from the Institute of Applied Psychology, The University of Stockholm, Stockholm, Sweden.

Isreal, J. B., Chesney, G. L., Wickens, C. D., and Donchin, E., 1980, P300 and tracking difficulty: Evidence for multiple resources in dual-task performance, Psychophysiology, 17:259-273.

Isreal, J. B., Wickens, C. D., Chesney, G. L., and Donchin, E., 1980, The event-related brain potential as an index of display-monitoring workload, Human Factors, 22:211-244.

Jex, H. R. and Clement, W. F., 1979, Defining and measuring perceptual-motor workload in manual control tasks: in: "Mental Workload: Its Theory and Measurement," N. Moray, ed., Plenum Press, New York.

Jiang, Q. and Beatty, J., 1981, Physiological assessment of operator workload during manual tracking, Proceedings of the 17th Annual Conference on Manual Control, Los Angeles, California.

Kelly, C. R. and Wargo, M. J., 1967, Cross-adaptive operator loading tasks, Human Factors, 9:395-404.

Knowles, W. B., 1963, Operator loading tasks, Human Factors, 5:151-161.

Lane, N. E., Streib, M. I., Glenn, F. A., and Wherry, R. J., 1981, The human operator simulator: An overview: in: "Manned Systems Design Methods, Equipment, and Applications," J. Morssel and K. F. Kraiss, eds., Plenum Press, New York.

Lane, N. E., Streib, M., and Wherry, R. J., Jr., 1977, The human operator simulator: Estimation of workload reserve using a simulated secondary task, Proceedings of the AGARD Conference on Methods to Assess Workload, No. AGARD-CP-216.

Linton, P. M., Jahns, D. W., and Chatelier, P. R., 1977, Operator workload assessment model: An evaluation of a VF/VA-V/STOL system, Proceedings of the AGARD Conference on Methods to Assess Workload, No. AGARD-CP-216.

Moray, N., ed., 1979, "Mental Workload: Its Theory and Measurement," Plenum Press, New York.

Moray, N., 1982, Subjective mental workload, Human Factors, 24:25-40.

Navon, D. and Gopher, D., 1977, On the economy of the human processing system, Psychological Review, 86:214-255.

North, R. A., 1977, Task components and demands as factors in dual-task performance, Report No. ARL-77-2/AFOSR-77-2, Aviation Research Laboratory, University of Illinois at Urbana-Champaign, Champaign, Illinois.

Notestine, J., 1983, Subjective workload assessment in a probability monitoring task and the effect of delayed ratings, Unpublished Master's Thesis, Wright State University, Dayton, Ohio.

Nygren, T. E., 1982, Conjoint measurement and conjoint scaling: A users guide, Technical Report No. AFAMRL-TR-82-22, U.S. Air Force Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio.

O'Donnell, R. D., 1979, Contributions of psychophysiological techniques to aircraft design and other operational problems, Report No. 244, NATO Advisory Group for Aerospace Research and Development.

Ogden, G. D., Levine, J. M., and Eisner, E. J., 1979, Measurement of workload by secondary tasks, Human Factors, 21:529-548.

Parks, D. L., 1979, Current workload methods and emerging challenges: in: "Mental Workload: Its Theory and Measurement," N. Moray, ed., Plenum Press, New York.

Quinn, T. J., Jauer, R. A., and Summers, P. I., 1982, Radar aided mission/aircrew capability exploration, RAM/ACE interim report--task II synthesis, Technical Report No. AFAMRL-TR-82-91, U.S. Air Force Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio.

Rahimi, M. and Wierwille, W. W., 1982, Evaluation of the sensitivity and intrusion of workload estimation techniques in piloting tasks emphasizing mediational activity, Proceedings of the 1982 IEEE International Conference on Cybernetics and Society, 593-597.

Reid, G. B., Eggemeier, F. T., and Nygren, T. E., 1982, An individual differences approach to SWAT scale development, Proceedings of the 1982 Human Factors Annual Meeting, 639-642.

Reid, G. B., Eggemeier, F. T., and Shingledecker, C. A., 1982, Subjective workload assessment technique, Proceedings of the 1982 AIAA Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics, 281-288.

Reid, G. B., Eggemeier, F. T., and Shingledecker, C. A., 1984, Workload analysis for the AMRAAM operational test and evaluation, Air Force Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio (in preparation).

Reid, G. B., Shingledecker, C. A., and Eggemeier, F. T., 1981a, Application of conjoint measurement to workload scale development, Proceedings of the 1981 Human Factors Society Annual Meeting, 522-526.

Reid, G. B., Shingledecker, C. A., Nygren, T. E., and Eggemeier, F. T., 1981b, Development of multidimensional subjective measures of workload, Proceedings of the 1981 IEEE International Confererence on Cybernetics and Society, 403-406.

Rolfe, J. M., 1971, The secondary task as a measure of mental load: in: "Measurement of Man at Work," W. T. Singleton, J. G. Fox, and D. Whitfield, eds., Taylor and Francis, London, 135-148.

Rolfe, J. M., 1976, The measurement of human response in man-vehicle control situations" in: "Monitoring Behavior and Supervisory Control," T. Sheridan and G. Johannsen, eds., Plenum Press, New York.

Sanders, A. F., 1979, Some remarks on mental load: in: "Mental Workload: Its Theory and Measurement," N. Moray, ed., Plenum Press, New York.

Schifflet, S. G., Linton, P. M., and Spicuzza, R. J., 1982, Evaluation of a pilot workload assessment device to test alternative display formats and control handling qualities, Proceedings of the 1982 AIAA Workshops on Flight Testing to Identify Pilot Workload and Pilot Dynamics, 222-233.

Shingledecker, C. A., 1980, Enhancing operator acceptance and noninterference in secondary task measures of workload, Proceedings of the Human Factors Society 1980 Annual Meeting, 674-677.

Shingledecker, C. A., 1983, Behavioral and subjective workload metrics for operational environments, Proceedings of the AGARD (AMP) Symposium, "Sustained Intensive Air Operations: Physiological and Performance Aspects," Paris, France (preprint).

Shingledecker, C. A., Acton, W. H., and Crabtree, M. S., 1983, Development and application of a criterion task set of workload metric evaluation, SAE Technical Paper Series, Paper No. 831419, Society of Automotive Engineers, Warrendale, Pennsylvania.

Shingledecker, C. A. and Crabtree, M. S., 1982, Subsidiary radio communications tasks for workload assessment in R&D circulations: II. Task sensitivity evaluation, Technical Report No. AFAMRL-TR-82-57, U.S. Air Force Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio.

Shingledecker, C. A., Crabtree, M. S., and Acton, W. H., 1982, Standardized tests for the evaluation and classification of workload metrics, Proceedings of the 1982 Human Factors Society Annual Meeting, 648-651.

Shingledecker, C. A., Crabtree, M. S., Simons, J. C., Courtright, J. F., and O'Donnell, R. D., 1980, Subsidiary radio communications tasks for workload assessment in R&D simulations: I. Task development and workload scaling, Technical Report No. AFAMRL-TR-80-126, U.S. Air Force Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio.

Skelly, J., Reid, G. B., and Wilson, G. R., 1983, B-52 full mission simulation: Subjective and physiological workload applications, Paper presented at the Second Aerospace Behavioral Engineering Technology Conference, Long Beach, California.

Wickens, C. D., 1981, Processing resources in attention, dual task performance, and workload assessment, Technical Report No. EPL-81-3, Engineering Psychology Research Laboratory, University of Illinois, Champaign, Illinois.

Wickens, C. D. and Derrick, W., 1981, Workload measurement and multiple resources, Proceedings of the 1981 IEEE Conference on Cybernetics and Society, 600-603.

Wickens, C. D., and Kessel, C., 1980, The processing resource demands of failure detection in dynamic systems, Journal of Experimental Psychology: Human Perception and Performance, 6:564-577.

Wierwille, W. W., 1979, Physiological measures of aircrew mental workload, Human Factors, 21:575-593.

Wierwille, W. W. and Casali, J. G., 1983(a), A validated rating scale for global mental workload measurement applications, Proceedings of 1983 Human Factors Society Annual Meeting, 129-133.

Wierwille, W. W. and Casali, J. G., 1983(b), The sensitivity and intrusion of mental workload estimation techniques in piloting tasks, IEOR Department Report No. 8309, Department of Industrial Engineering and Operations Research, Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

Wierwille, W. W. and Connor, S. A., 1983, Evaluation of 20 workload measures using a psychomotor task in a moving-base aircraft simulator, Human Factors, 25:1-16.

Wierwille, W. W. and Williges, R. C., 1978, Survey and analysis of operator workload assessment techniques, Report No. S-78-101, Systemetrics, Inc., Blacksburg, Virginia.

Williges, R. C. and Wierwille, W. W., 1979, Behavioral measures of aircrew mental workload, Human Factors, 21:549-574.

Zipoy, D. R., Premselaar, S. J., Gargett, R. E., Belyea, I. L., and Hall, J. J., 1970, Integrated information presentation and control systems study, Vol. I, System development concepts, Air Force Flight Dynamics Laboratory, Technical Report No. AFFDL-TR-70-79, Wright-Patterson Air Force Base, Ohio.

PAPER PRESENTED AT THE AMERICAN PSYCHOLOGICAL ASSOCIATION MTG, AUGUST 1982.

A CONCEPTUAL FRAMEWORK FOR DEVELOPMENT OF A WORKLOAD ASSESSMENT METHODOLOGY

F. Thomas Eggemeier
Wright State University
and
Systems Research Laboratories, Inc.

and

Robert D. O'Donnell
USAF Aerospace Medical Research Laboratory

ABSTRACT

Based on a review of the current literature, a conceptual framework which incorporates major elements related to operator workload has been developed. The framework treats workload as a multidimensional construct with important physiological, subjective, and behavioral components. An important implication of the framework is that, at present, a comprehensive workload assessment methodology should include a number of measures, including subjective, physiological, and performance-based metrics.

INTRODUCTION

A primary concern of human factors engineering during system development and evaluation is to assure that the performance demands imposed by a system do not exceed the human operator's capacity to process information. Mental workoad is the term which has been used in referring to the degree or percentage of the operator's information processing capacity which is expended in meeting system demands. The growing complexity of modern systems and associated increases in workload have increased the likelihood of exceeding or approaching the limitations of an operator's processing capacity. As a consequence, the need for reliable and sensitive methods of assessing the workload imposed by alternative design options has assumed increased importance in recent years.

In order to address the need for viable workload metrics, a large number of individual measurement techniques have been developed and documented in a number of reviews (e.g., Gartner and Murphy, 1976; Williges and Wierwille, 1979; Moray, 1982). Williges and Wierwille (1979), for example, identified 28 specific workload assessment techniques which have been proposed in the recent literature. Despite the fact that there are numerous individual techniques, most workload measures can generally be classified as belonging to one of the three major categories:

1. Subjective measures, such as rating scales, which require the operator to rate or somehow characterize the subjective workload associated with performance of a particular task or with a system design option.

2. Psychophysiological measures, such as heart rate variability, which derive an index of workload from some aspect of the operator's physiological response to task or system demands.

3. Performance-based measures, such as secondary task methodology, which use some aspect of the operator's capability to perform within the system as a measure of workload.

All three classes of measures have been used extensively during the past 10 years with varying degrees of success. Given the large number of measures, a major issue in developing a workload assessment methodology is choosing those measures that should be incorporated into the methodology. Unfortunately, although there is considerable agreement in the literature regarding the importance of workload, there has not been substantial agreement concerning the most appropriate means by which to assess it. Much of the failure to agree on the appropriate assessment technique stems from the fact that workload is a multidimensional concept (e.g., Johannsen, Moray, Pew, Rasmussen, Sanders, and Wickens, 1979; Sanders, 1979; Williges and Wierwille, 1979; Sheridan and Simpson, 1979; White, 1971) which has been used in referring to several different aspects of system or task demand, operator effort and information processing capacities, operator performance, and systems performance (e.g., Sheridan and Stassen, 1979).

In order to provide a means for organizing the numerous factors associated with workload, a general conceptual framework which addresses major elements of the workload construct has been developed. The major objective in assembling the framework was to help guide the development of a comprehensive workload assessment methodology which would tap essential elements of the workload construct. The framework is largely based on current conceptualizations of workload and theories regarding the nature of capacity limitations within the human information processing system. The purpose of this paper is to provide an overview of the conceptual framework and to discuss implications of the framework for workload assessment.

CONCEPTUAL FRAMEWORK FOR WORKLOAD

Despite the diversity in emphasis among various approaches to workload, three general elements which represent major factors in most current theoretical treatments can be identified. These general elements include:

1. Some characterization of the demands placed on the operator by the system or task.

2. Some expression of the capacity or effort expenditure required by the operator to deal with the demands.

3. The level of operator performance that results from the interaction of task demands and capacity/effort expenditure.

Figure 1 illustrates these major elements and their interrelationship at a very general level.

Insert Figure 1 here

As indicated in Figure 1, system demands are imposed on the operator and must be related to or mapped onto the operator's processing capacity or resources. Depending upon the efficiency of the operator in mapping demands onto available capacities and the degree of effort expended by the operator, a particular level of performance results. Effort is intended to represent a construct similar to that described by Kahneman (1973) and Jahns (1973), and represents a nonspecific input which is required to activate information processing structures. Most conceptualizations of workload include each of these elements, and a variety of workload assessment procedures have been developed to evaluate aspects of each element.

Several current models and theoretical statements concerning workload (e.g., Jahns, 1973; Welford, 1978; Sanders, 1979; Sheridan and Simpson, 1979; Johannsen et al., 1979) treat major elements outlined in Figure 1 in more detail. These models and statements generally maintain that workload is clearly a multidimensional construct that reflects the interaction of such elements as task and system demands, operator processing capacities and effort, subjective performance criteria, operator information processing strategies, and operator training or experience.

In addition to suggesting that workload itself is multidimensional, a number of investigators have also maintained that several elements of workload depicted in Figure 1 are themselves multidimensional.

Jahns (1973), for instance, categorized sources of input load or system demand into three classes: (1) environmental, (2) situational, and (3) procedural. Environmental demands included factors such as temperature, humidity, noise, and acceleration which can serve as sources of load for an operator. Demands that were characterized as situational were composed of elements such as display and control characteristics, display-control arrangement, vehicle dynamics, and crewstation volume. These types of demand reflect traditional human factors engineering concerns, and play a central role in determining the overall level of input load experienced by an operator. Finally, procedural demands included elements such as mission or task duration, standard system operating procedures, the mission itself, and briefings/instructions given to the operator. Jahns does not specify how the various sources of demand combine to determine overall levels of load, but it is clear that input load is considered to be multidimensional and is described as a vector rather than a scalar quantity.

A similar conclusion has been reached by Johannsen et al. (1979) with respect to the concept of effort (Kahneman, 1973; Jahns, 1973). Johannsen et al. (1979) pointed out that the notion of effort is extremely complex and can be interpreted in several different ways. One sense of effort identified by Johannsen et al. is related to the physiological activation as measured by a number of indices (e.g., muscle tension, respiration rate) that occurs when an operator is exposed to progressive increases in mental load. A second related sense of effort is the subjective feeling experienced by an operator under a high load condition. Johannsen et al. assume that such feelings represent the products of muscular tension and changes in physiological variables such as blood pressure and heart rate. It is noted that an operator may actually feel loaded and effortful despite the fact that there is no change in the adequacy of performance. This can be explained by assuming that increases in task difficulty lead to a period during which the operator is working harder in order to preclude performance decrements. The clear implication is that effort has both physiological and subjective dimensions, each of which may reflect increases in load prior to any actual decrements in operator performance.

Operator processing capacity which is depicted in Figure 1 has also been characterized as multidimensional in some recent descriptions of the information processing system. This position assumes that the information processing system may be described as a series of internal processing structures, each with its own processing capacity or resources which are not exchangeable with any other structure. Current theoretical positions that are consistent with this model are the multiple resources or structure-specific resource models of processing capacity (e.g., Wickens, 1979, 1981; Navon and Gopher, 1979; Sanders, 1979; North, 1977; Kantowitz and Knight, 1976). According to this theory, a considerable amount of capacity may remain unused in responding to a particular task's demands, because ony a limited number of processing structures may be involved in that response. An important consequence of this position identified by Sanders (1979), Wickens (1979), and Gopher (1978), is that mental load cannot be conceptualized as a single dimension; and ultimately, a task may have to be described in terms of multidimensional patterns of mental load. A central question in workload specification becomes one of determining the extent to which various processing resources are involved in a task which, in turn, specifies the pattern of mental load. Several theorists (e.g., North, 1977; Sanders, 1979) have suggested candidate dimensions of resources that could be used in describing the resource demand composition of a task. The most comprehensive position, however, is that of Wickens (1981) who, on the basis of available evidence, has identified three primary dichotomous dimensions that appear to define separate resources. These dimensions include:

1. Stages of information processing (perceptual/central processing operations versus response selection and execution).

2. Modalities of perception (auditory versus visual).

3. Codes of information processing and response (spatial-manual versus verbal-vocal).

Under the system proposed by Wickens, an adequate description of the load imposed by a task should include specification of its demand composition on each of the three dimensions. The implication of the multiple resources theory is clear: the capacity to process information is multidimensional and characterizations of the load imposed by a task should reflect that multidimensionality.

The general conceptual framework outlined previously can now be modified to be more descriptive of elements related to workload. Figure 2 illustrates the modified conceptual framework. In the modified framework, workload and operator performance continue to represent the product of an interaction of several factors. In accordance with Jahns' position, system demands are represented as multidimensional, including the classes of environmental, situational, and procedural load. Effort is also conceptualized as reflecting several dimensions, including strong physiological and subjective components. Likewise, operator processing resources have been revised to reflect several of the dimensions represented in the multiple resources view of the information processing system.

Insert Figure 2 here

The major conclusion that follows from the present framework is that workload, as currently conceptualized, represents a multidimensional construct that includes important physiological, subjective, and behavioral components. The multidimensionality is reflected not only in the interaction of several elements to determine levels of load, but also in the multidimensionality of several elements themselves. The multidimensional nature of the framework has important implications for workload assessment, and these are discussed in the next section.

IMPLICATIONS OF THE CONCEPTUAL FRAMEWORK FOR WORKLOAD ASSESSMENT

A major implication of the current framework is that no single measurement technique will provide a comprehensive means for assessment of load. Since current theory generally maintains that there are important physiological, subjective, and behavioral components of load, a comprehensive approach to assessment of workload should include physiological, subjective, and behavioral or performance-based measures. At present, it is not clear how such measures might ultimately be combined to provide a multidimensional index of load, but it should be clear from the foregoing discussion that subjective and physiological measures can potentially provide information not afforded by performance-based measures, and vice versa. Therefore, it appears that the most viable approach to comprehensive workload assessment would be a battery of measures, including subjective, physiological, and performance-based components which could be applied to derive indices of several components of load. Similar conclusions have been drawn recently by several others. Johannsen et al. (1979), for example, concluded that workload contains behavioral, performance, physiological, and subjective components and indicated that appropriate measures would be required for each component. Williges and Wierwille (1979), in their review of behavioral workload assessment procedures, maintain that due to the multi-dimensionality of workload, it appears unlikely that any single measure will be completely sufficient for characterizing load. Williges and Wierwille conclude that multiple measures, includng dimensions of subjective opinion, spare mental capacity, primary tasks, and physiological correlates need to be considered. Similarly, Jahns (1973) noted that because of the complex interactions involved in determining levels of operator effort, a broad spectrum of measurement techniques for operator workload need to be investigated. White (1971) also indicated that work-load is multidimensional and, at present, cannot be defined adequately in terms of any single measure.

Further support for the necessity of considering multiple measures of load comes from the fact that when several measures of workload are applied in a situation, they commonly exhibit some degree of dissociation (e.g., Borg, 1978; Hicks and Wierwille, 1979; Dornic and Andersson, 1980; Wickens and Derrick, 1981; and Moray, 1982). This type of result can, of course, be interpreted within the multidimensional workload framework outlined above, since different measurement techniques can be assumed to be maximally sensitive to different dimensions of load. Wickens and Derrick (1981), for instance, have noted that a lack of correspondence among workload measures can be attributed to the multidimensional nature of information processing resources that are assumed to underlie performance and workload. Based on the results of several experiments in which dissociation occurred between physiological, primary task, subjective, and secondary task measures, Wickens and Derrick proposed that some measures may be considered more generally sensitive to overall levels of demand for resources anywhere within the information processing system, while others are more diagnostic in the sense of specifically reflecting demands imposed on particular resources (e.g., percpetual versus motor output) within the processing system. It was suggested that some physiological measures such as heart rate variability might be sensitive to the total demand placed on all resources of the system, even if the demand was imposed such that no single resource was overloaded and no performance decrement occurred. Subjectve measures were also viewed as generally sensitive to demands imposed anywhere in the system, while secondary task measures were thought to be more diagnostic in providing information regarding the specific resources demanded by a task.

Wickens and Derrick concluded that since individual measures might provide different types of information about load, choice of measure might be dictated in part by the purpose of taking the measure. If the total load imposed by a task or system were to be assessed, some physiological or subjective techniques might be used in conjunction with primary task performance measures. However, if the specific locus of an overload (e.g., central processing versus response execution) were to be identified in order to provide more diagnostic information to a human factors engineer, then secondary task methodology might be the more appropriate choice. Note that in this type of scheme, several measurement techniques could be used in a complementary fashion, with some subjective or physiological measures providing an initial index of overall task or system workload, and secondary tasks or other physiological measures (e.g., evoked cortical potential) being used subsequently to develop more precise information regarding the locus of specific overloads.

The clear implication that follows from the conceptual framework and from the noted dissociation of workload measures is that, at present, a comprehensive workload assessment methodology will require physiological, subjective, and performance-based measurement techniques. The major goals for research in the workload metric development area suggested by the framework are to: [1] initially identify the most sensitive measure(s) within each category of assessment technique; and [2] conduct a systematic comparison of the information provided by each category of assessment technique so that optimal combination(s) of measures required for a comprehensive workload assessment methodology can ultimately be established.

REFERENCES

Borg, G. Subjective aspects of physical and mental load. Ergonomics, 1978, 21, 215-220.

Dornic, S. and Andersson, O. Difficulty and effort: A perceptual approach. Stockholm, Sweden: Reports from the Department of Psychology, the University of Stockholm, Report No. 566, November 1980.

Gartner, W. B. and Murphy, M. R. Pilot workload and fatigue: A critical survey of concepts and assessment techniques. Washington, D.C.: National Aeronautics and Space Administration, Report No. NASA-TN-D-8365, November 1976.

Hicks, T. G. and Wierwille, W. W. Comparison of five mental workload assessment procedures in a moving-base driving simulator. Human Factors, 1979, 21, 129-143.

Jahns, D. W. A concept of operator workload in manual vehicle operations. Meckenheim, Germany: Forschungsinstitut fur Anthropotechnik, Report No. 14, 1973.

Johannsen, G., Moray, N., Pew, R., Rasmussen, J., Sanders, A., and Wickens, C. Final report of the experimental psychology group. In: N. Moray (ed.), Mental Workload: Its theory and measurement. New York: Plenum Press, 1979.

Kahneman, D. Attention and effort. Englewood Cliffs, New Jersey: Prentice-Hall, 1973.

Kantowitz, B. H. and Knight, J. L. Testing tapping timesharing, II: Auditory secondary task. Acta Psychologica, 1976, 40, 342-362.

Moray, N. Subjective mental workload. Human Factors, 1982, 24, 25-40.

Navon, D. and Gopher, D. On the economy of the human processing system, Psychological Review, 1979, 86(3), 214-255.

North, R. A. Task components and task demands as factors in dual-task performance. University of Illinois, Aviation Research Laboratory, Report No. ARL-77-2, January 1977.

Sanders, A. F. Some remarks on mental load. In: N. Moray (ed.), Mental workload: Its theory and measurement. New York: Plenum Press, 1979.

Sheridan, T. B. and Simpson, R. W. Toward the definition and measurement of the mental workload of transport pilots. Cambridge, Massachusetts: Massachusetts Institute of Technology Flight Transportation Laboratory Report, FTL Report R79-4, January 1979.

Sheridan, T. B. and Stassen, H. G. Definitions, models and measures of human workload. In: N. Moray (ed.), Mental workload: Its theory and measurement, New York: Plenum Press, 1979.

Welford, A. T. Mental workload as a function of demand, capacity, strategy, and skill. Ergonomics, 1978, 21, 151-167.

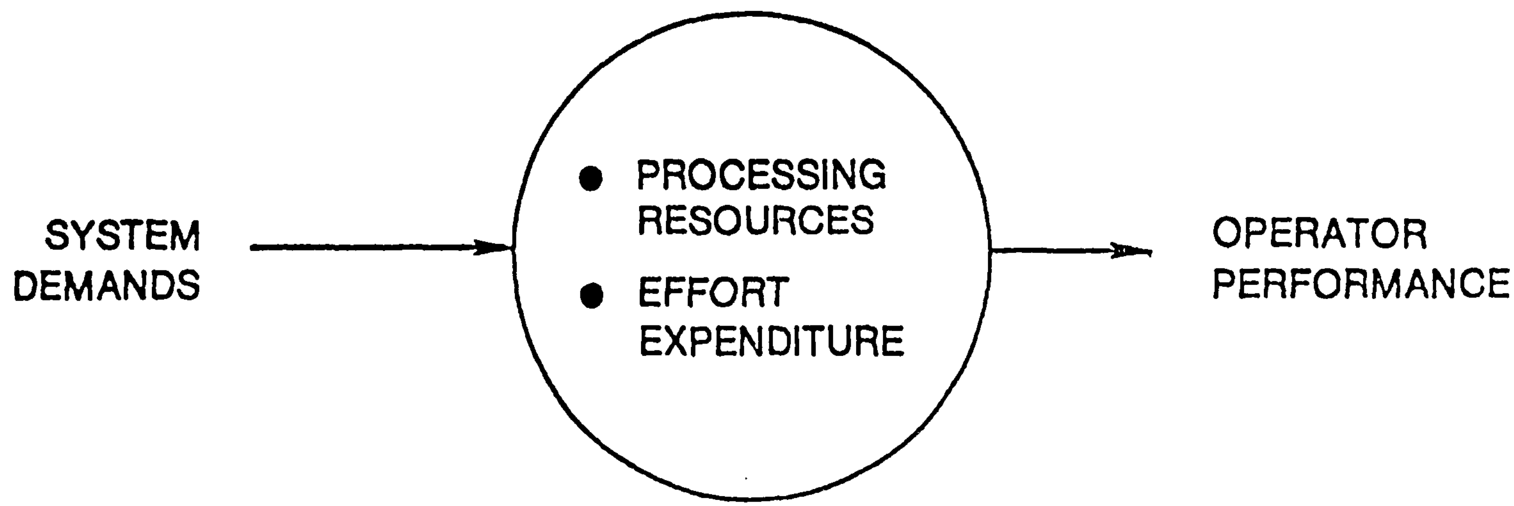
White, R. T. Task analysis methods: Review and development of techniques for analyzing mental workoad in multiple task situations. St. Louis, Missouri: McDonnell-Douglas Corporation, MDC J 5291, September 1971.

Wickens, C. D. Measures of workload, stress, and secondary tasks. In N. Moray (ed.), Mental workload: Its theory and measurement. New York: Plenum Press, 1979.

Wickens, C. D. Processing resources in attention, dual task performance, and workload assessment. Engineering Psychology Research Laboratory Technical Report EPL-81-3, University of Illinois, 1981.

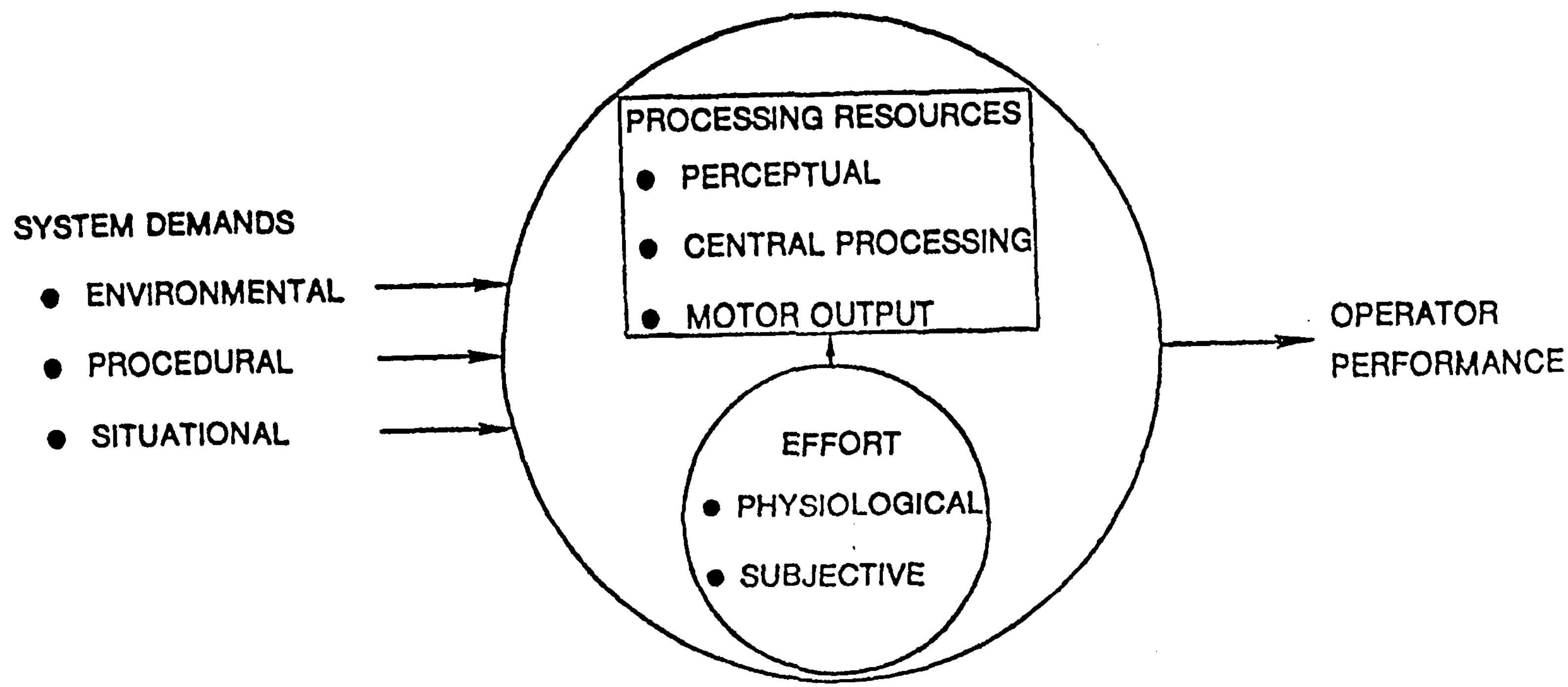
Wickens, C. D. and Derrick, W. Workload measurement and multiple resources. In: Proceedings of the 1981 IEEE Conference on Cybernetics and Society, 1981, 600-603.

Williges, R. C. and Wierwille, W. W. Behavioral measures of aircrew mental workload. Human Factors, 1979, 21, 549-574.



GENERAL CONCEPTUAL FRAMEWORK FOR WORKLOAD

FIGURE 1



MODIFIED CONCEPTUAL FRAMEWORK FOR WORKLOAD

FIGURE 2

WORKLOAD ASSESSMENT METHODOLOGY

- SATISFACTORY OPERATOR PERFORMANCE REQUIRES THAT DEMAND NOT EXCEED PROCESSING CAPACITY
- WORKLOAD IS THE PORTION OF OPERATOR CAPACITY EXPENDED IN MEETING SYSTEM DEMANDS
- ASSESSMENT OF WORKLOAD A CENTRAL CONCERN THROUGHOUT SYSTEM DEVELOPMENT

HE-83-10-1

WORKLOAD ASSESSMENT METHODOLOGY

- MAJOR CLASSES OF WORKLOAD ASSESSMENT TECHNIQUES
  - — SUBJECTIVE
  - — PERFORMANCE BASED
  - — PHYSIOLOGICAL

HE-83-10-2

WORKLOAD ASSESSMENT METHODOLOGY

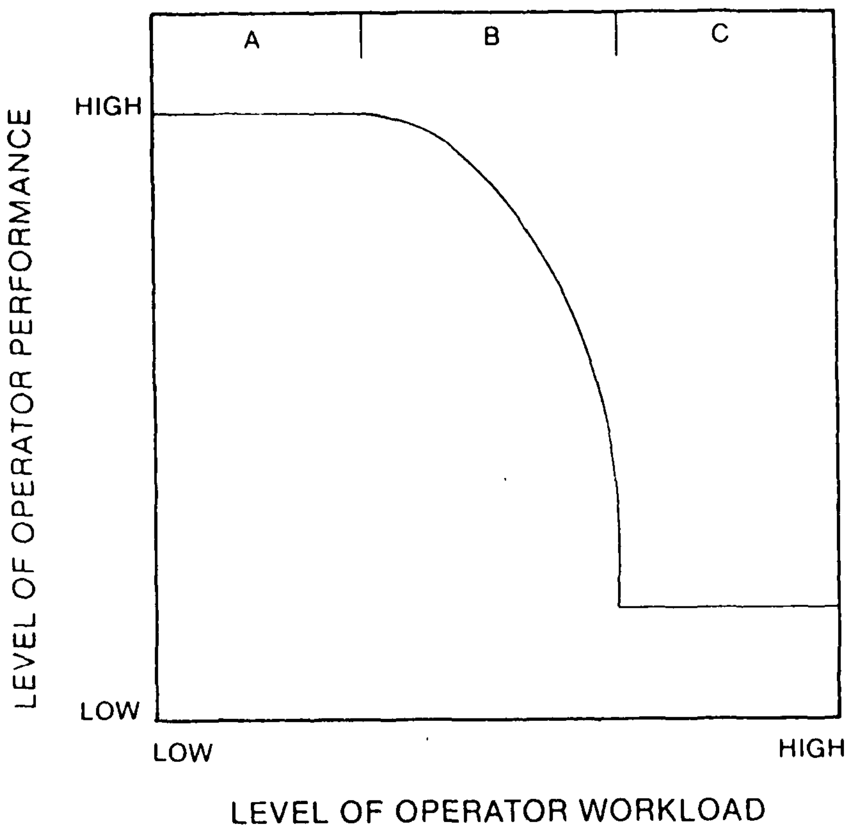
- WORKLOAD METRIC SELECTION CRITERIA
  - — SENSITIVITY
  - — DIAGNOSTICITY
  - — DEGREE OF INTRUSIVENESS
  - — EASE OF IMPLEMENTATION
  - — OPERATOR ACCEPTANCE

HE-83-10-3

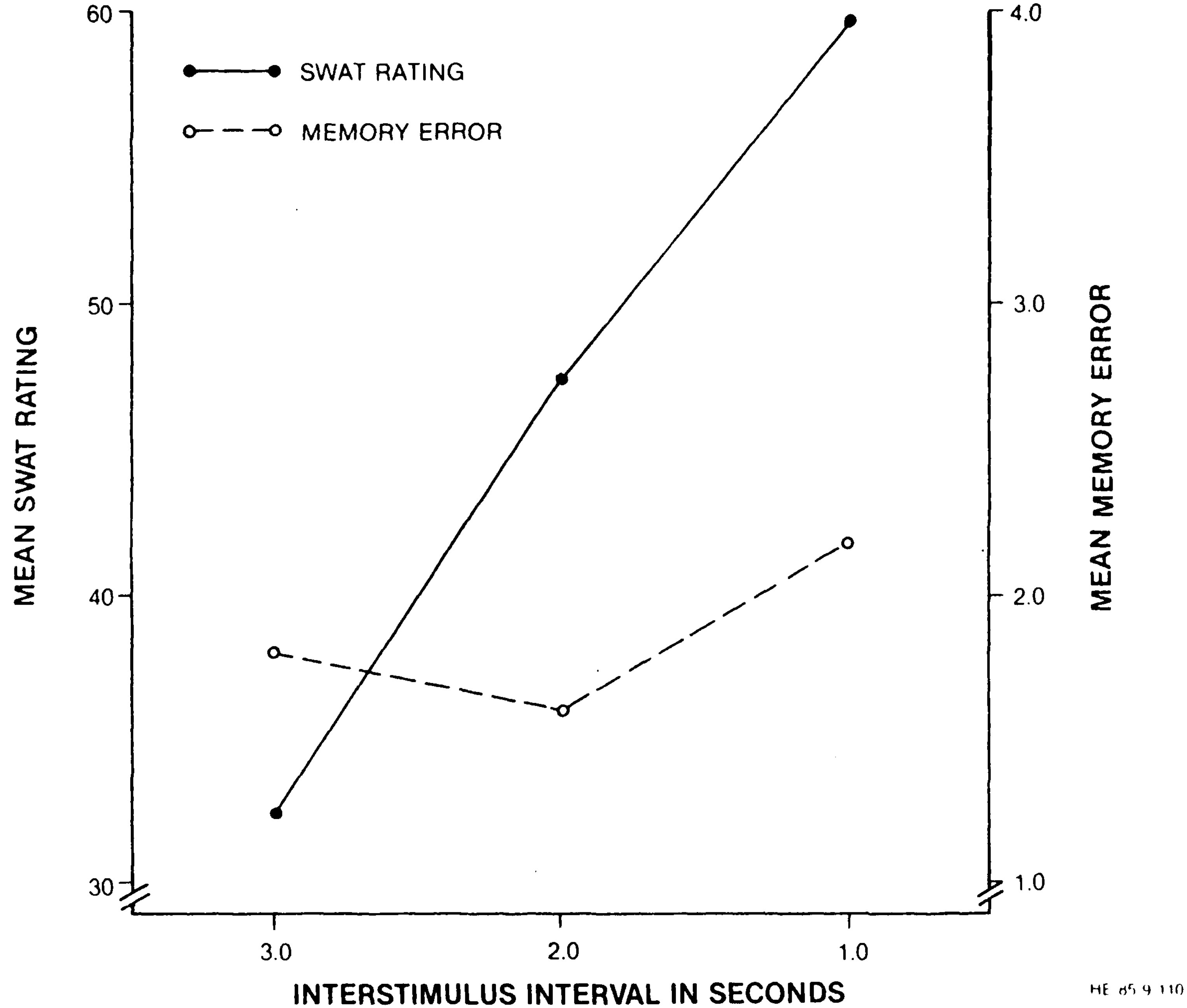
WORKLOAD ASSESSMENT METHODOLOGY

- SENSITIVITY
  - — DEGREE OF SENSITIVITY RELATED TO NATURE OF QUESTION
  - — PRIMARY TASK MEASURES DISTINGUISH OVERLOAD FROM NON OVERLOAD
  - — OTHER TYPES OF TECHNIQUES GENERALLY MORE SENSITIVE

HE-83-10-4


A
B
C
HIGH
LOW
LEVEL OF OPERATOR PERFORMANCE
LOW
HIGH
LEVEL OF OPERATOR WORKLOAD

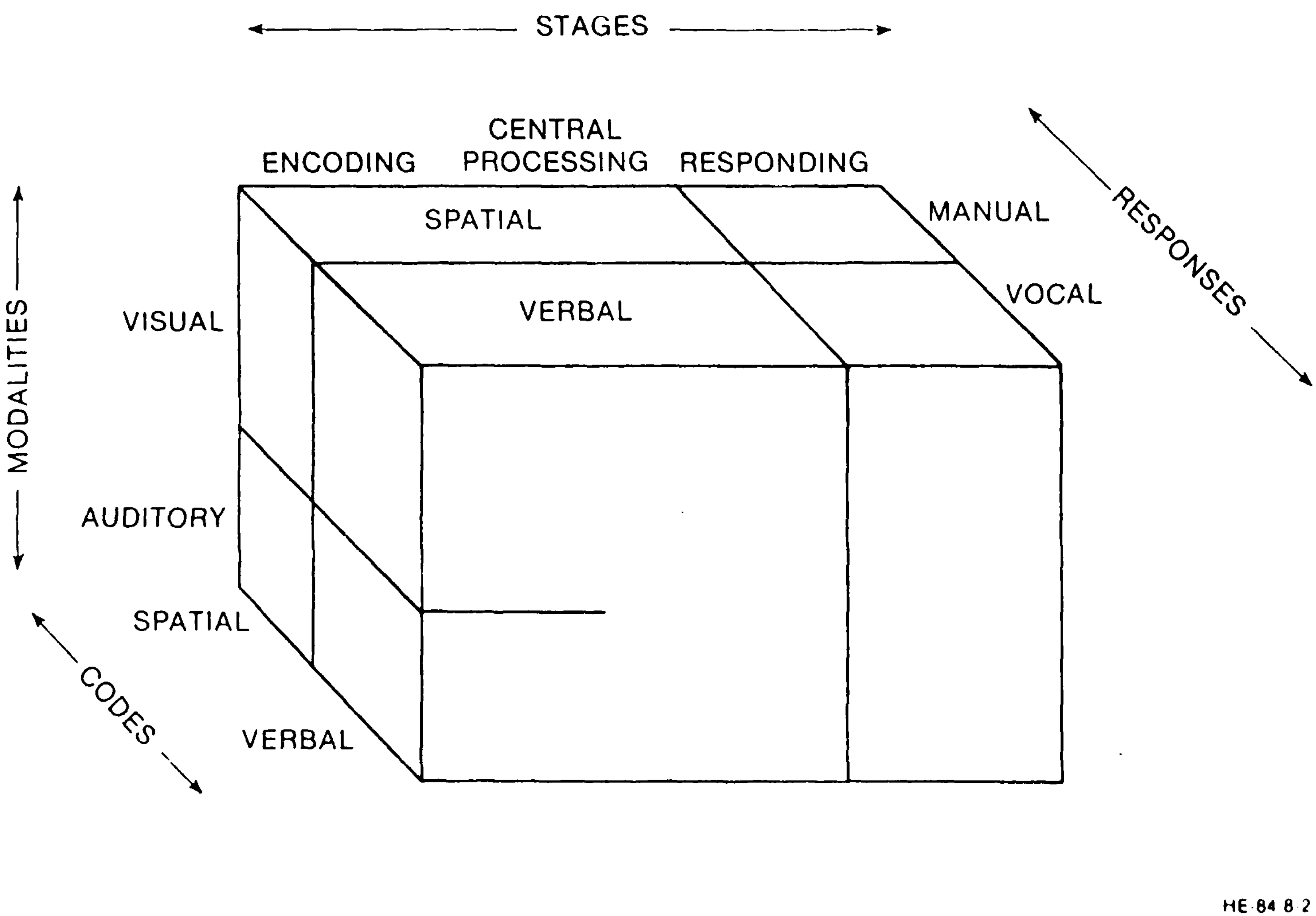
HE-84-8-3


SWAT RATING
MEMORY ERROR
MEAN SWAT RATING
60
50
40
30
MEAN MEMORY ERROR
4.0
3.0
2.0
1.0
3.0
2.0
1.0
INTERSTIMULUS INTERVAL IN SECONDS
HE 85 9 110

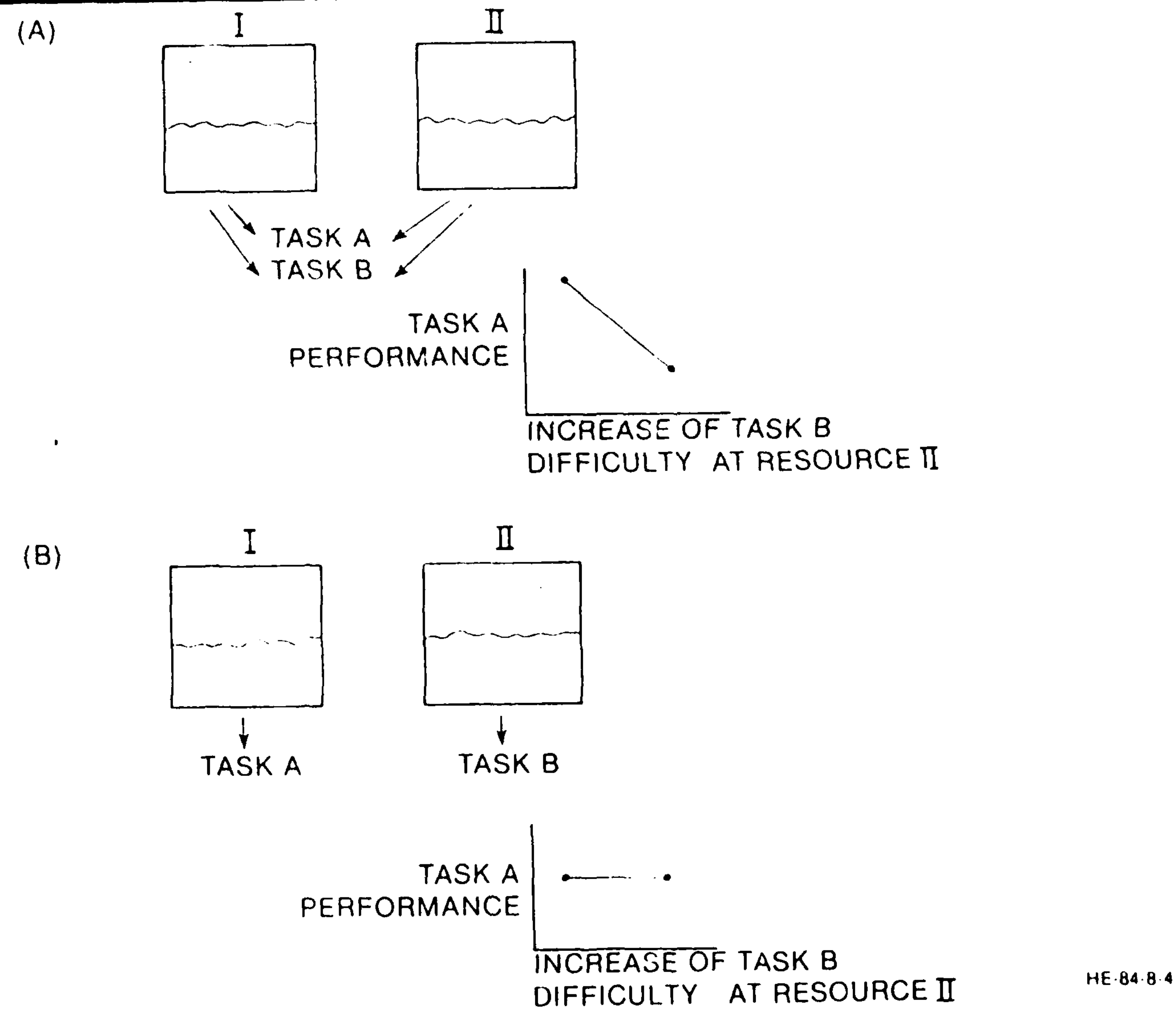
WORKLOAD ASSESSMENT METHODOLOGY

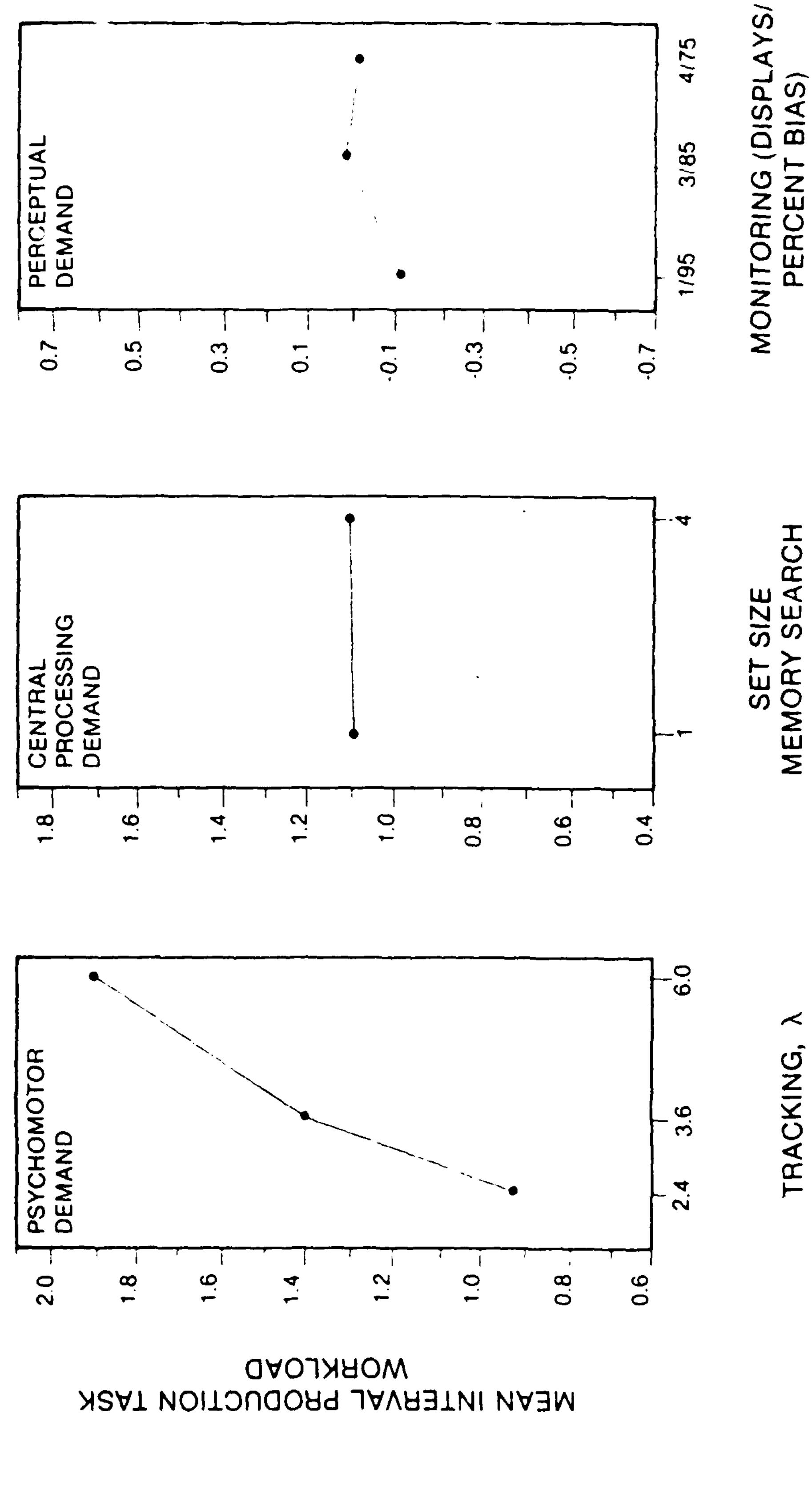
- DIAGNOSTICITY
  - — CAPABILITY TO DISCRIMINATE DIFFERENCES IN SPECIFIC RESOURCE EXPENDITURE
  - — DEGREE OF DIAGNOSTICITY RELATED TO NATURE OF QUESTION
  - — SECONDARY TASK METHODOLOGY AND SOME PHYSIOLOGICAL MEASURES (e.g. EVOKED CORTICAL POTENTIAL) CONSIDERED HIGHLY DIAGNOSTIC
  - — SUBJECTIVE TECHNIQUES AND PRIMARY TASK MEASURES CONSIDERED LESS DIAGNOSTIC

HE-83-10-5


STAGES
ENCODING
CENTRAL PROCESSING
RESPONDING
SPATIAL
VERBAL
MANUAL
VOCAL
RESPONSES
MODALITIES
VISUAL
AUDITORY
SPATIAL
VERBAL
CODES
HE-84 8 2

RESOURCES


(A)
I
II
TASK A
TASK B
TASK A PERFORMANCE
INCREASE OF TASK B DIFFICULTY AT RESOURCE II
(B)
I
II
TASK A
TASK B
TASK A PERFORMANCE
INCREASE OF TASK B DIFFICULTY AT RESOURCE II
HE-84-8-4


MEAN INTERVAL PRODUCTION TASK
WORKLOAD
PSYCHOMOTOR
DEMAND
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
2.4
3.6
6.0
TRACKING, λ
CENTRAL
PROCESSING
DEMAND
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
1
4
SET SIZE
MEMORY SEARCH
PERCEPTUAL
DEMAND
0.7
0.5
0.3
0.1
-0.1
-0.3
-0.5
-0.7
1/95
3/85
4/75
MONITORING (DISPLAYS/
PERCENT BIAS)

HE 84-8-5

WORKLOAD ASSESSMENT METHODOLOGY

- DEGREE OF INTRUSIVENESS
  - — INTRUSION UNDESIRABLE FOR BOTH PRACTICAL AND THEORETICAL REASONS
  - — INTRUSION HAS BEEN A PROBLEM WITH MANY SECONDARY TASK APPLICATIONS
  - — SUBJECTIVE, PHYSIOLOGICAL, AND PRIMARY TASK METRICS MINIMIZE POTENTIAL INTRUSION

HE-83-10-6

WORKLOAD ASSESSMENT METHODOLOGY

- IMPLEMENTATION REQUIREMENTS
  - SUBJECTIVE TECHNIQUES MINIMIZE INSTRUMENTATION REQUIREMENTS
  - SECONDARY TASK AND SOME SUBJECTIVE TECHNIQUES CAN REQUIRE OPERATOR TRAINING
  - TRADITIONAL APPLICATIONS OF ALL EMPIRICAL TECHNIQUES REQUIRE LABORATORY/SIMULATION FACILITIES OR OPERATIONAL EQUIPMENT

HE-83-10-7

WORKLOAD ASSESSMENT METHODOLOGY

- OPERATOR ACCEPTANCE
  - NO SYSTEMATIC DATA BASE
  - INFORMAL EVIDENCE SUGGESTS HIGH ACCEPTANCE OF SUBJECTIVE TECHNIQUES AND PRIMARY TASK MEASURES
  - SOME POTENTIAL FOR LIMITATIONS IN ACCEPTANCE OF SECONDARY TASK AND PHYSIOLOGICAL PROCEDURES

HE-83-10-8

WORKLOAD ASSESSMENT METHODOLOGY

- NO ASSESSMENT TECHNIQUE SATISFIES ALL APPLICABLE CRITERIA
- WORKLOAD ASSESSMENT REQUIRES COMPLEMENTARY USE OF SEVERAL METRICS
  - — SUBJECTIVE
  - — PERFORMANCE-BASED
  - — PHYSIOLOGICAL

HE-83-10-9

WORKLOAD ASSESSMENT METHODOLOGY

- CAPABILITY OF METRICS TO MEET CRITERIA SUGGEST SOME GUIDELINES FOR USAGE

- TYPE OF QUESTION TO BE ANSWERED AND PRACTICAL CONSTRAINTS CRITICAL IN CHOICE OF METRIC

HE-83-10-10

SUMMARY OF WORKLOAD ASSESSMENT TECHNIQUE CAPABILITIES

	SENSITIVITY	DIAGNOSTICITY	INTRUSIVENESS	IMPLEMENTATION REQUIREMENTS	OPERATOR ACCEPTANCE
PRIMARY TASK MEASURES	DISCRIMINATE OVERLOAD FROM NONOVERLOAD SITUATIONS. USED TO DETERMINE IF OPERATOR PERFORMANCE WILL BE ACCEPTABLE WITH A PARTICULAR DESIGN OPTION	NOT CONSIDERED DIAGNOSTIC. REPRESENTS A GLOBAL MEASURE OF WORKLOAD THAT IS SENSITIVE TO OVERLOADS ANYWHERE WITHIN THE OPERATOR'S PROCESSING SYSTEM.	NONINTRUSIVE SINCE NO ADDITIONAL OPERATOR PERFORMANCE OR REPORT REQUIRED.	INSTRUMENTATION FOR DATA COLLECTION CAN RESTRICT USE IN OPERATIONAL ENVIRONMENTS USE REQUIRES MOCKUPS, SIMULATORS, OR OPERATIONAL EQUIPMENT. IMPOSES LIMITS ON USE DURING EARLY SYSTEM DEVELOPMENT NO OPERATOR TRAINING REQUIRED	NO SYSTEMATIC DATA. NO REASON TO EXPECT NEGATIVE OPERATOR OPINION
SECONDARY TASK METHODS	CAPABLE OF DISCRIMINATING LEVELS OF CAPACITY EXPENDITURE IN NONOVERLOAD SITUATIONS. USED TO ASSESS *RESERVE* CAPACITY AFFORDED BY A PRIMARY TASK. CAN BE USED TO ASSESS THE POTENTIAL FOR OVERLOAD AMONG DESIGN OPTIONS	CAPABLE OF DISCRIMINATING SOME DIFFERENCES IN RESOURCE EXPENDITURE (E.G., *CENTRAL PROCESSING VERSUS* MOTOR). DIAGNOSTICITY SUGGESTS COMPLEMENTARY USE WITH MORE GENERALLLY SENSITIVE MEASURES, WITH THE LATTER INITIALLY IDENTIFYING OVERLOADS AND SECONDARY TASKS BEING USED SUBSEQUENTLY TO PINPOINT THE LOCUS OF OVERLOAD	PRIMARY TASK INTRUSION HAS REPRESENTED A PROBLEM IN MANY APPLICATIONS, PARTICU*LARLY IN THE LABORATORY*. DATA ARE NOT EXTENSIVE IN OPERATIONAL ENVIRONMENTS. SEVERAL TECHNIQUES (E.G., EMBEDDED SECONDARY TASK, ADAPTIVE PROCEDURES) HAVE BEEN DESIGNED TO CONTROL INTRUSION. *POTENTIAL FOR INTRUSION* COULD LIMIT USE IN OPERATIONAL ENVIRONMENTS	INSTRUMENTATION FOR DATA COLLECTION CAN RESTRICT USE IN OPERATIONAL ENVIRONMENTS. BUT SOME TASKS HAVE BEEN INSTRUMENTED FOR IN-FLIGHT USE. USE REQUIRES MOCKUPS, SIMULATORS, OR OPERATIONAL EQUIPMENT. IMPOSES LIMITS ON USE DURING EARLY SYSTEM DEVELOPMENT SOME OPERATOR TRAINING USUALLY REQUIRED TO STABILIZE SECONDARY TASK PERFORMANCE	NO SYSTEMATIC DATA. REQUIREMENT TO PERFORM SECONDARY TASK COULD DISTRACT OPERATOR. TECHNIQUE SUCH AS EMBEDDED SECONDARY TASK SHOULD MINIMIZE ANY ACCEPTANCE PROBLEMS
PHYSIOLOGICAL TECHNIQUES	CAPABLE OF DISCRIMINATING LEVELS OF CAPACITY EXPENDITURE IN NONOVERLOAD SITUATIONS. CAN BE USED TO ASSESS THE RELATIVE POTENTIAL FOR OVERLOAD AMONG DESIGN OPTIONS	SOME TECHNIQUES (E.G., EVENT-RELATED BRAIN POTENTIAL) APPEAR DIAGNOSTIC OF SOME RESOURCES, WHILE OTHER MEASURES (E.G., PUPIL DIAMETER) APPEAR MORE GENERALLY SENSITIVE. CHOICE OF TECHNIQUE DEPENDENT ON PURPOSE OF MEASUREMENT (SCREENING FOR ANY OVERLOAD VERSUS IDENTIFYING LOCUS OF OVERLOAD)	INTRUSION DOES NOT APPEAR TO REPRESENT A MAJOR PROBLEM, ALTHOUGH THERE ARE DATA TO INDICATE THAT SOME INTERFERENCE CAN OCCUR.	INSTRUMENTATION FOR DATA COLLECTION CAN RESTRICT USE IN OPERATIONAL ENVIRONMENTS. USE REQUIRES MOCKUPS, SIMULATORS, OR OPERATIONAL EQUIPMENT. IMPOSES LIMITS ON USE DURING EARLY SYSTEM DEVELOPMENT NO OPERATOR TRAINING REQUIRED	NO SYSTEMATIC DATA. INSTRUMENTATION AND RECORDING EQUIPMENT COULD REPRESENT POTENTIAL PROBLEM, BUT NO SIGNIFICANT PROBLEMS REPORTED IN LITERATURE
SUBJECTIVE TECHNIQUES	CAPABLE OF DISCRIMINATING LEVELS OF CAPACITY EXPENDITURE IN NONOVERLOAD SITUATIONS. CAN BE USED TO ASSESS THE RELATIVE POTENTIAL FOR OVERLOAD AMONG DESIGN OPTIONS	NOT CONSIDERED DIAGNOSTIC. AVAILABLE EVIDENCE INDICATES THAT RATING SCALES REPRESENT A GLOBAL MEASURE OF LOAD. LACK OF DIAGNOSTICITY SUGGETS USE AS A GENERAL SCREENING DEVICE TO DETERMINE IF OVERLOAD EXISTS ANYWHERE WITHIN TASK PERFORMANCE	INTRUSION DOES NOT APPEAR TO REPRESENT A SIGNIFICANT PROBLEM. MOST APPLICATIONS REQUIRE RATING SCALE COMPLETION SUBSEQUENT TO TASK PERFORMANCE AND, THEREFORE, PRESENT NO INTRUSION PROBLEM	INSTRUMENTATION REQUIRED IS USUALLY MINIMAL, PERMITTING USE IN A NUMBER OF ENVIRONMENTS. TRADITIONAL APPLICATIONS REQUIRE MOCKUPS, SIMULATORS, OR OPERATIONAL EQUIPMENT. IMPOSES LIMITS ON USE DURING EARLY SYSTEM DEVELOPMENT. RECENT PROJECTIVE USE PROVIDES POTENTIAL FOR APPLICATION DURING EARLY STAGES. SOME FAMILIARIZATION WITH PROCEDURES CAN BE REQUIRED	NO SYSTEMATIC DATA. INFORMAL EVIDENCE SUGGESTS THAT SEVERAL RATING SCALES ENJOY A HIGH DEGREE OF OPERATOR ACCEPTANCE

HE-84-8-6

SESSION 1. SUBJECTIVE TECHNIQUES

Presenters

G. Reid: Subjective Workload Assessment Technique (SWAT)

S. Hart: Bi-Polar Subjective Technique

Subjective Workload Assessment Technique (SWAT)
by G. Reid

SWAT CARD SORT INSTRUCTIONS FOR SUBJECTS

During the course of this experiment, you will be asked to quantify the mental workload required to complete the tasks you will be performing. Mental Workload refers to how hard you work to accomplish some task, groups of tasks, or an entire job. The workload imposed on you at any one time consists of a combination of various dimensions which contribute to the subjective feeling of workload. The Subjective Workload Assessment Technique (SWAT) defines these dimensions as 1) Time Load, 2) Mental Effort Load, and 3) Psychological Stress Load.

For the purpose of SWAT, the three dimensions have been assigned three levels. The dimensions and their levels are described in the following paragraphs.

TIME LOAD

Time load refers to the fraction of the total time that you are busy. When time load is low, sufficient time is available to complete all of your mental work with some time to spare. As time load increases, spare time drops out and some aspects of performance overlap and interrupt one another. This overlap and interruption can come from performing more than one task or from different aspects of performing the same task. At higher levels of time load, several aspects of performance often occur simultaneously, you are busy, and interruptions are very frequent.

Time load may be rated on the three point scale below.

(1) Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.
(2) Occasionally have spare time. Interruptions or overlap among activities occur frequently.
(3) Almost never have spare time. Interruptions or overlap among activities are very frequent, or occur all the time.

MENTAL EFFORT LOAD

As described above time load refers to the amount of time one has available to perform a task or tasks. In contrast, mental effort load is an index of the amount of attention or mental effort required by a task regardless of the number of tasks to be performed or any time limitations. When mental effort load is low, the concentration and attention required by a task is minimal and performance is nearly automatic. As the demand for mental effort increases, due to task complexity or the amount of information which must be dealt with in order to perform adequately, the degree of concentration and attention required increases. High mental effort load demands total attention or concentration due to task complexity or the amount of information that must be processed.

Mental effort load may be rated using the three point scale below.

(1) Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.

(2) Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required.

(3) Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

PSYCHOLOGICAL STRESS LOAD

Stress load refers to the contribution to total workload of any conditions that produce anxiety, frustration or confusion while performing a task or tasks. At low levels of stress, one feels relatively relaxed. As stress increases, confusion, anxiety or frustration increase and greater concentration and determination are required to maintain control of the situation.

Psychological stress load may be rated on the three point scale below.

(1) Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.
(2) Moderate stress due to confusion, frustration, or anxiety noticeably adds to workload. Significant compensation is required to maintain adequate performance.
(3) High to very intense stress due to confusion, frustration, or anxiety. High to extreme determination and self-control required.

Each of the three dimensons just described contribute to workload during performance of a task or group of tasks. Note that although all three factors may be correlated, they need not be. For example, one can have many tasks to perform in the time available (high time load) but the tasks may require little concentration (low mental effort.) Likewise, one can be anxious and frustrated (high stress) and have plenty of spare time between relatively simple tasks. Since the three dimensons contributing to workload are not necessarily correlated, please treat each dimension individually and give independent assessments of the time load, mental effort load, and stress load that you experience in performing the following tasks.

One of the most important features of SWAT is its unique scoring system. SWAT uses a procedure to find separate scoring weights for each level of a dimensions. Then, it determines a distinctive workload scale for each person. This scaling system greatly improves the precision of the workload ratings you will give later.

In order to develop your individual scale, we need information from you regarding the amount of workload you feel is imposed by various combinations of the dimensions described above. We get this information by having you rank order the workload associated with each of the combinations.

In order for you to rank order the workload for each of the combinations, you have been given a set of 27 cards with the combinations from each of the three dimensions. Each card contains a different combination of levels of Time Load, Mental Effort, and Psychological Stress. Your job is to sort the cards so that they are rank ordered according to the level of workload represented on each.

In completing your card sorts, please consider the workload imposed on a person by the combination represented in each card. Arrange the cards from the lowest workload condition through the highest condition. You may use any strategy you choose in rank ordering the cards. One strategy that proves useful is to arrange the cards into a number of preliminary stacks representing "High", "Moderate", and "Low" workload. Individual cards can be exchanged between stacks, if necessary, and then rank ordered within stacks. Stacks can then be recombined and checked to be sure that they represent your ranking of lowest to highest workload. However, the choice of strategy is up to you and you should choose the one that works best for you.

There is no "school solution" to this problem. There is no correct order. The correct order is what, in your judgment best describes the progression of workload from lowest to highest for a general case rather than any specific event. That judgment differs for each of us. The letters you see on the back of the cards are to allow us to arrange the cards in a previously randomized sequence so that everyone gets the same order. If you examine your deck you will see the order on the back runs from A through Z and then ZZ.

Please remember:

1) The card sort is being done so a workload scale may be developed for you. This scale will have a distinct workload value for each possible combinations of Time Load, Mental Effort Load and Psychological Stress Load.

Time	Effort	Stress	Workload Scale
			0
1	1	1	
.	.	.	
.	.	.	
.	.	.	
.	.	.	
.	.	.	
.	.	.	
.	.	.	
3	3	3	100

2) When performing the card sorts, use the descriptors printed on the cards. Please remember not to sort the cards based on a particular task (such as flying an airplane). Sort the cards according to your general view of workload and how important you consider the dimensions of time, mental effort, and psychological stress load to be.

3) During the actual experiment, you will accomplish the desired task. Then, you will provide a SWAT score based on your opinion of the mental workload required to perform the task. This SWAT score will consist of one number from each of the three dimensions.

For example, a possible SWAT score is 1 - 2 - 2. This represents a 1 for Time Load, a 2 for Mental Effort Load, and a 2 for Psychological Stress Load.

4) We are not asking for your preference concerning Time, Mental Effort and Psychological Stress Load. Some people may prefer to be "busy" rather than "idle" in either the Time Load, Mental Effort load, or Psychological Stress Load dimension. We are not concerned with this preference. We need information on how the three dimensions and the three levels of each one will affect the level of workload as you see it. You may prefer a 2 - 2 - 2 situation instead of a 1 - 1 - 1 situation. But, you should still realize that the 1 - 1 - 1 situation imposes less workload on you and leaves a greater reserve capacity.

From this point until you have completed the sorting will probably take 30 minutes to an hour.

Please feel free to ask questions at any time. Thank you for your cooperation.

Almost never have spare time. Interruptions or overlap among activities are very frequent, or occur all the time.

Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required.

Moderate stress due to confusion, frustration, or anxiety noticably adds to workload. Significant compensation is required to maintain adequate performance.

Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.

Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.

Moderate stress due to confusion, frustration, or anxiety noticably adds to workload. Significant compensation is required to maintain adequate performance.

Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.

Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required.

High to very intense stress due to confusion, frustration, or anxiety. High to extreme determination and self-control required.

Almost never have spare time. Interruptions or overlap among activities are very frequent, or occur all the time.

Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required.

Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.

Occasionally have spare time. Interruptions or overlap among activites occur frequently.

Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.

High to very intense stress due to confusion, frustration, or anxiety. High to extreme determination and self-control required.

Occasionally have spare time. Interruptions or overlap among activites occur frequently.

Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required.

High to very intense stress due to confusion, frustration, or anxiety. High to extreme determination and self-control required.

Occasionally have spare time. Interruptions or overlap among activites occur frequently.

Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required.

Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.

Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.

Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.

High to very intense stress due to confusion, frustration, or anxiety. High to extreme determination and self-control required.

Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.

Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.

Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.

Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

Moderate stress due to confusion, frustration, or anxiety noticably adds to workload. Significant compensation is required to maintain adequate performance.

Almost never have spare time. Interruptions or overlap among activities are very frequent, or occur all the time.

Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

Moderate stress due to confusion, frustration, or anxiety noticably adds to workload. Significant compensation is required to maintain adequate performance.

Occasionally have spare time. Interruptions or overlap among activites occur frequently.

Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.

Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.

Almost never have spare time. Interruptions or overlap among activities are very frequent, or occur all the time.

Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.

Moderate stress due to confusion, frustration, or anxiety noticably adds to workload. Significant compensation is required to maintain adequate performance.

Occasionally have spare time. Interruptions or overlap among activites occur frequently.

Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required.

Moderate stress due to confusion, frustration, or anxiety noticably adds to workload. Significant compensation is required to maintain adequate performance.

Occasionally have spare time. Interruptions or overlap among activites occur frequently.

Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

High to very intense stress due to confusion, frustration, or anxiety. High to extreme determination and self-control required.

Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.

Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

High to very intense stress due to confusion, frustration, or anxiety. High to extreme determination and self-control required.

Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.

Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.

Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.

Almost never have spare time. Interruptions or overlap among activities are very frequent, or occur all the time.

Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required.

High to very intense stress due to confusion, frustration, or anxiety. High to extreme determination and self-control required.

Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.

Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required.

Moderate stress due to confusion, frustration, or anxiety noticably adds to workload. Significant compensation is required to maintain adequate performance.

Occasionally have spare time. Interruptions or overlap among activites occur frequently.

Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.

Almost never have spare time. Interruptions or overlap among activities are very frequent, or occur all the time.

Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.

Occasionally have spare time. Interruptions or overlap among activites occur frequently.

Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.

Moderate stress due to confusion, frustration, or anxiety noticably adds to workload. Significant compensation is required to maintain adequate performance.

Almost never have spare time. Interruptions or overlap among activities are very frequent, or occur all the time.

Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.

Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.

Almost never have spare time. Interruptions or overlap among activities are very frequent, or occur all the time.

Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

High to very intense stress due to confusion, frustration, or anxiety. High to extreme determination and self-control required.

Almost never have spare time. Interruptions or overlap among activities are very frequent, or occur all the time.

Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.

High to very intense stress due to confusion, frustration, or anxiety. High to extreme determination and self-control required.

Occasionally have spare time. Interruptions or overlap among activites occur frequently.

Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

Moderate stress due to confusion, frustration, or anxiety noticably adds to workload. Significant compensation is required to maintain adequate performance.

Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.

Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required.

Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.

INFORMATION FOR PROGRAM OPERATION

This appendix describes the procedures for inputting and analyzing SWAT scale development data. The program is interactive and the user menus or screens are designed to help guide you through the analysis process. The interface has been designed for ease of use and should facilitate data analysis. As an additional aid the <ESC> key has been programmed to be the user's "panic button". At anytime while you are working with the SWAT program that you become confused or lost in the analysis, you can press the <ESC> key and return to the main menu and all data will be retained by the program. In this way you can always get back to a place where you know your way out and avoid losing data.

This appendix is organized for a complete description of each of the screens with which the user will interact. First, a sample screen will be presented for reference, and then the screen functions will be described. To the degree possible, the order of execution for a routine analysis will be followed. We will deviate from this practice only to completely describe all of the options presented on a given screen. Output information resulting from this analysis includes prototype correlations, axiom test results, and the conjoint scaling solution.

MAIN MENU

As can be seen in Figure 1, this screen has three main functions. The first is to allow entering of date, study name, and additional comments for labeling a data set. The second function is to specify the name of the file to be used to later identify the data set and the third function is to select the program options.

When you assign the file name on this screen it becomes the file the program uses when the data set is stored to disk. The file name should be different from the study name or some variant of the study name to allow for multiple analysis you may do within a study. As always, if you want to retrieve a data file for something, the file name you use must be exactly correct. For example, you may want to access a previous data file to modify it by adding additional subjects.

Upon entering a new file name, the data entry screen, which is described next, will appear. If you select a name of an existing file:

a. You will be asked to verify that you wish to use data from a previous file. This is a safeguard to prevent accidently writing over an existing data file.

b. The screen will change to show the updated version of the Main Menu screen including the information just entered by you as depicted in Figure 2. You will be asked to choose one of the following options to direct the program to specific options described below:

OPTIONS

1. Edit comments- This option allows you to change any or all of the comments on the first screen.

2. Data entry- This option takes you to the data entry screen to add, subtract, or otherwise modify data.

3. Program setup- Selection of this option takes you to the program setup screen where you will select options that permit you to specify the analysis you want performed. See the description of Program setup screen for more details.

DATA ENTRY SCREEN:

Upon selection of "Data Entry" on the main menu screen or when you enter a new file name, a formatted screen as depicted in Figure 3 will be displayed that allows you to enter or modify data. The number of subjects that was previously entered on the main menu will be displayed at the top of the screen, and sufficient space for data for this number of subjects will be created. The following options then exist:

1. Enter Data - If you select this option the cursor automatically moves to the position for subject #1, card 1-1-1 and data may be entered. The rank assigned to card 1-1-1 by this subject should be entered and then the cursor will move to subject #2. The program is set up to accept the rank assigned by all subjects for card 1-1-1 before the cursor automatically goes to the position for card 1-1-2 for subject #1. In this manner the program will step through all 27 cards for the specified number of subjects. You may manually override this sequence and move the cursor to any position on the data entry screen in the manner that any screen editor is used. Every third row on this screen is highlighted to make it easier to keep track of where you are as you scroll through the data. As you reach the end of the displayed data, the screen will scroll to the next position to allow continuation of data entry. You may also use the arrow keys to scroll to any position on this screen.

2. Save Data - When you select this option the program saves the data currently on the data entry screen to a diskette or hard disk file with the name you previously specified on the main menu screen. When the data are saved, this screen changes to that of Figure 4 and the program displays the following options:

a. Print data - This allows you to obtain a hard copy of the input data set for easy reference. Since the entire data table cannot fit on the screen at one time, this may be a more desirable way to proof and edit your input data.

b. Edit data - This option allows you to position the cursor at any point on the data entry screen and make changes to the data.

c. Program setup - This option directs the program to proceed to the next screen in order to continue the normal analysis process.

3. Edit data - If you are working with a data set which was previously entered, this function allows you to return to the data entry table and make changes to the data. Use the arrow keys to move the cursor to the appropriate position so that you may make the desired changes. Pressing <Return> or using the arrow keys to move the curser to the next position will enter the data into the computer's memory.

4. Escape - This option takes you back to the Main Menu and easily allows you to leave the program, help you get reoriented if you were lost or confused, or make changes in earlier screens. When this option is chosen, there will be a prompt to ask whether you want to save any data entered on the screen to a diskette file.

REMEMBER, data is saved ONLY through the SAVE DATA function, or saving during an ESCAPE.

PROGRAM SETUP SCREEN:

This screen is presented in Figure 5 and gives you options for the types of analyses to be performed. Enter the number(s) of the analyses which you want performed. There are three main analyses:

1. Prototype correlations and Kendall's Coefficient of Concordance
2. Axiom Testing
3. Scaling Solution

1. The Prototype Correlations analysis performs a Spearman's rank order correlation on each of the subject's rank order data with the six possible "perfect" rankings. The pattern of correlations indicates the importance a subject places on the three dimensions of the SWAT definition of workload. The Kendall's Coefficient of Concordance is an index of the degree of agreement among the group of subjects about the order of the 27 cards. A high Kendall's indicates substantial agreement about the order and therefore about the relative importance of the three dimensions.

2. The Axiom testing section performs the axiom tests for independence, joint independence, and double cancellation. This is done to check for violations of these axioms which may invalidate the additive model as being a suitable model to use for the conjoint scaling routine. Only a summary of

the three axiom test is automatically displayed, but the complete history of results for all the tests may be viewed or printed by selection of the appropriate option.

3. Scaling solution - Produces the scale values that result from the conjoint scaling routine. These values are then used as the workload scores in subsequent analyses of results from the Event Scoring Phase.

There are three methods for handling the scale development data: group solution, prototyped solutions, and individual solutions. The determination of which solution is best is based on study objectives and the results of the correlations, Kendall's Coefficient, and axiom tests. For a more detailed explanation of interpreting these analyses, refer to the section entitled "Interpretation of Output". If prototype or individual axiom tests or scaling solutions are chosen, the program will ask for specific prototype groups and/or individual subjects you desire to include in the analysis. This option is useful if, for example, you wish to run three subjects individually. Their subject number may be specified and the program will only include their analysis. If, for example, you specify a prototype solution by separating the subjects into prototype groups, the program will use the suggested prototype groups provided in the next screen to be described. [If you wish to change the prototype, press <Return> to procede with the analysis.]

Upon pressing <Return>, the program will proceed with the analyses previously chosen on this Program Setup screen. Note that due to the criticality of the prototype correlations section, it will ALWAYS be performed, even if it was not selected. Additionally, the program always performs the three sections in the order described above, regardless of the order of the options in your selection. While the program is executing, the following message will appear: "PROGRAM EXECUTING".

PROTOTYPE ANALYSIS:

This screen is presented in Figure 6 and displays the Spearman's Rho correlations for each subject's data with the six "perfect" prototype rankings, the suggested prototype groupings based on these correlations, and the Kendall's Coefficient of Concordance for the agreement among the entire group of subjects in this analysis.

There are several options for this screen:

1. Print - This option sends the correlation matrix, prototypes, and Kendall's Coefficient to a printer.

2. Change prototypes - This option may be desired to group the subjects slightly differently based on the experimenter's

evaluation of the pattern of correlations. To leave a particular subject out of the analysis, simply change that subject's suggested prototype to an "L". The program will ignore this subject's data. This option can be used in a case where a subject belongs to a particular prototype group, but his/her data also contains a large number of axiom test violations, indicating an unacceptable amount of error in the data. In this situation it may be desirable to exclude this subject from the scaling solution, or attempt to remove the ambiguity in the data through additional information obtained from the subject. To end any changes to the prototypes, simply push F1 to return to the rest of the options.

8. Return to Program Setup - This option allows you to return to the program setup screen and choose additional or different analyses. This option is used in two situations:

a. If prototype correlations was the only option previously chosen in the program setup. This allows you to choose additional analyses based on the results of the prototype correlations and Kendall's coeffienent of concordance.

b. If, based on the results of the prototype correlations and Kendall's coefficient of concordance, you decide to chose different configurations for axiom testing and scaling solutions. For example, say you originally choose prototype correlations, group axioms, and group scale. Now you observe that the subjects are not homogeneous, with a Kendal's of .72. You may wish to simply proceed with separating the subject population into prototype groups rather than a single group solution. Choosing (3), Return to Program Setup, will cancel the original options and allow you to enter new options.

4. Proceed to Next Option Chosen In Program Setup - This option displays the results of the next option previously chosen, either axiom tests or scaling solution.

5. Escape - This option returns you to the Main Menu in case you run into problems and need to start over.

SUMMARY OF AXIOM TEST VIOLATIONS:

When option 4 of the previous screen (Procede to Next Option Chosen in Program Setup) is chosen, and you had previously chosen axiom tests as an option in the program setup screen, the screen shown in Figure 7 will be displayed. This screen presents a summary of the violations for the simple independence axiom tests, which is the most critical axiom for acceptance of an additive model. The screen heading will state whether the summary is for a single group, a prototype group, or an individual subject. If the summary is for the entire group of subjects, the following options exist:

1. Print summary - This will print a hardcopy of just the summary of the independence axioms.
2. Print all axioms - This will print a hardcopy of the results of the entire set of axiom tests, including independence, joint independence, and double cancellation.
3. View all axioms - Allows you to view on screen the results of the entire set of axiom tests. To scroll through all of the results, use the <UP ARROW> and <DOWN ARROW> keys. When finished viewing the entire results, the <--> key will return you to this option screen.
4. Proceed To Next Option Chosen In Program Setup - This option will display the scaling solution for the group of subjects, if scaling was chosen in program setup. If scaling solution was not chosen the program will go back to the Program Setup screen.
5. Return To Program Setup - As in the previous screen, this option allows you to return to the program setup and change the configuration of the analysis. Other options previously chosen will be cancelled, and new ones can be selected.
6. Escape - As before, this returns you to the main menu in case of problems or for a fast way to exit the program.

If the Axiom Test Summary is for a prototype group, the following options exist:

1. Print summary - same as for group.
2. Print all asioms - same as for group.
3. View all axioms - same as for group.
4. Proceed to next option chosen in program setup - will display the scaling solution for this same prototype group, if scaling was chosen in the program setup screen.
5. Return to program setup - same as for group.
6. Go to next prototype - will display summary information of axioms violations for the next prototype chosen in the program setup. Since the prototype is a result of the correlation matrix, the program is aware of the next prototype group.

If the summary is for an individual, the following options exist:

1. Print summary - same as for a group.
2. Print all axioms - same as for a group.
3. View all axioms - same as for a group.
4. Proceed To Next Option Chosen In Program Setup - same as for a group.
5. Return to Program Setup - same as for a group.
6. Go to Next Individual - Will display summary information of independence axiom violations for the next individual chosen in the program setup.

In all three cases, when "Proceed To Next Option Chosen" is selected on the Summary of Axiom Violations screen and a scaling solution option had been chosen on the Program Setup Screen or upon selection of the "Scaling" option on the Main Menu screen, the program will display the following information on the Scaling Information Screen:

1. The last five iterations of the Theta and Tau values for the scaling solution. For a more detailed explanation of these values, refer to the section "Interpretation of Output".
2. The rescaled values for each level of the subscales. These are the additive values which, through all possible combinations, form the 27 values of the scaling solution.
3. The approximate relative importance of each factor. This indicates the amount of change from level 1 of a dimension to level 3 of the same dimension.

******* COMMENTS AND MAIN MENU *******

TODAY'S DATE:
(mmddyy)

STUDY NAME: FILE NAME:

NUMBER OF SUBJECTS:

COMMENT:
COMMENT:
COMMENT:

** USE A SEPARATE DATA DISK FOR EACH STUDY **

```
*******  COMMENTS AND MAIN MENU  *******

      TODAY'S DATE:
                  (mmddyy)

      STUDY NAME:                          FILE NAME:work.dat


      NUMBER OF SUBJECTS:12

   COMMENT:
   COMMENT:
   COMMENT:


                              MAIN
                              MENU

                1  -EDIT COMMENTS            4  -CHANGE PARAMETERS
                2  -DATA ENTRY               5  -END THE PROGRAM
                3  -PROGRAM SETUP

                        MAKE A SELECTION:
```

```
******* COMMENTS AND MAIN MENU *******

        TODAY'S DATE:
                       (mmddyy)

        STUDY NAME:                              FILE NAME:work.dat

        NUMBER OF SUBJECTS:12

COMMENT:
COMMENT:
COMMENT:


                                    WARNING
                           FILE work.dat EXISTS
                           Work with the existing file (y/n) ?  y

** USE A SEPARATE DATA DISK FOR EACH STUDY **
```

*** ENTER SUBJECT DATA IN THIS TABLE ***

12 SUBJECTS

F1 -SAVE DATA
F2 -EDIT DATA
F3 -ENTER DATA
F4 -PRINT DATA
ESC -MAIN MENU

CARD	1	2	3	4	5	6	7	8
111 N	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
112 B	4.00	2.00	4.00	5.00	4.00	5.00	10.00	4.00
113 W	8.00	13.00	7.00	8.00	15.00	17.00	19.00	7.00
121 F	3.00	3.00	2.00	2.00	3.00	3.00	4.00	3.00
122 J	6.00	6.00	9.00	11.00	19.00	10.00	12.00	8.00
123 C	13.00	17.00	14.00	19.00	23.00	20.00	20.00	15.00
131 X	9.00	10.00	3.00	6.00	7.00	7.00	5.00	6.00
132 S	16.00	18.00	16.00	16.00	14.00	14.00	15.00	16.00
133 M	21.00	21.00	21.00	23.00	24.00	21.00	21.00	23.00
211 U	2.00	4.00	5.00	3.00	2.00	2.00	2.00	2.00
212 G	7.00	5.00	8.00	9.00	5.00	6.00	11.00	9.00
213 Z	12.00	16.00	19.00	13.00	8.00	19.00	22.00	17.00
221 V	5.00	8.00	10.00	4.00	9.00	4.00	3.00	10.00
222 Q	11.00	9.00	11.00	15.00	13.00	11.00	13.00	11.00
223 ZZ	18.00	19.00	13.00	20.00	17.00	22.00	23.00	20.00
231 K	14.00	11.00	20.00	12.00	10.00	9.00	6.00	14.00
232 E	19.00	20.00	18.00	18.00	21.00	16.00	14.00	19.00
233 R	25.00	23.00	25.00	21.00	25.00	23.00	24.00	26.00

```
******* PROGRAM SETUP *******

TO RUN ANY OF THESE PROGRAMS OR COMBINATIONS  OF PROGRAMS
   CHOOSE THE CORRESPONDING NUMBER(S) AND PRESS   RETURN

          1 - PROTOTYPE CORRELATIONS AND KENDALL'S

          2 - GROUP AXIOMS

          3 - GROUP SCALE

          4 - PROTOTYPE AXIOMS

          5 - PROTOTYPE SCALE

          6 - INDIVIDUAL AXIOMS

          7 - INDIVIDUAL SCALES

        ESC - MAIN MENU

            OPTIONS CHOOSEN:
```

```
       *  PROTOTYPE ANALYSIS OF EACH SUBJECTS DATA  *

  THE KENDALL'S COEFFICIENT OF CONCORDANCE WAS: W =

      SPEARMAN RANK CORRELATION (RS) FOR EACH SUBJECT

                                             SUGGESTED      F1   - CHANGE PROTOTYPE
SUB. #    TES   TSE   ETS   EST   SET   STE  PROTOTYPE      F2   - PRINT
                                                            F3   - RETURN TO PROGRAM
                                                                   SETUP
                                                            F4   - GO TO NEXT OPTION
                                                                   CHOOSEN IN
                                                                   PROGRAM SETUP
                                                            ESC  - MAIN MENU


            WORKING...
```

* PROTOTYPE ANALYSIS OF EACH SUBJECTS DATA *

THE KENDALL'S COEFFICIENT OF CONCORDANCE WAS: W = .8344

SPEARMAN RANK CORRELATION (RS) FOR EACH SUBJECT

PRESS F1 TO QUIT

SUB. #	TES	TSE	ETS	EST	SET	STE	SUGGESTED PROTOTYPE
1	.80	.78	.79	.77	.71	.71	T
2	.69	.68	.77	.79	.77	.74	E
3	.74	.74	.71	.70	.71	.72	T
4	.66	.69	.67	.70	.78	.78	S
5	.50	.54	.64	.73	.86	.81	S
6	.58	.67	.55	.63	.90	.91	S
7	.46	.51	.56	.64	.79	.76	S
8	.70	.72	.75	.78	.82	.81	S
9	.29	.43	.41	.58	.99	.96	S
10	.49	.54	.64	.74	.90	.85	S
11	.63	.63	.73	.77	.78	.74	S
12	.43	.30	.96	1.00	.60	.43	E

```
*******  SUMMARY OF AXIOM VIOLATIONS  *******

                    GROUP ANALYSIS

         INDEPENDENCE
T INDEPENDENT OF E AND S =  15. FAILURES OUT OF 84 TESTS
E INDEPENDENT OF T AND S =  12. FAILURES OUT OF 78 TESTS
S INDEPENDENT OF T AND E =   0. FAILURES OUT OF 78 TESTS
         DOUBLE CANCELATION
DOUBLE CANCELLATION IN T x E =   0. FAILURES OUT OF  0 TESTS
DOUBLE CANCELLATION IN E x S =   0. FAILURES OUT OF  0 TESTS
DOUBLE CANCELLATION IN S x T =   0. FAILURES OUT OF  4 TESTS
         JOINT INDEPENDENCE
T x E INDEPENDENT OF S =   3. FAILURES OUT OF 78 TESTS
E x S INDEPENDENT OF T =  14. FAILURES OUT OF 84 TESTS
S x T INDEPENDENT OF E =   8. FAILURES OUT OF 78 TESTS


                 OPTIONS - GROUP
           F1   - CONTINUE RUN
           F2   - PRINT SUMMARY OF AXIOM VIOLATIONS
           F3   - PRINT COMPLETE AXIOM HISTORY
           ESC  - MAIN MENU
```

```
*******  SUMMARY OF AXIOM VIOLATIONS  *******

                 PROTOTYPE ANALYSIS
                   TIME PROTOTYPE

        INDEPENDENCE
T INDEPENDENT OF E AND S =  17. FAILURES OUT OF 84 TESTS
E INDEPENDENT OF T AND S =  22. FAILURES OUT OF 78 TESTS
S INDEPENDENT OF T AND E =   0. FAILURES OUT OF 78 TESTS
        DOUBLE CANCELATION
DOUBLE CANCELLATION IN T x E =   0. FAILURES OUT OF  0 TESTS
DOUBLE CANCELLATION IN E x S =   0. FAILURES OUT OF  2 TESTS
DOUBLE CANCELLATION IN S x T =   0. FAILURES OUT OF  2 TESTS
        JOINT INDEPENDENCE
T x E INDEPENDENT OF S =  10. FAILURES OUT OF 78 TESTS
E x S INDEPENDENT OF T =  27. FAILURES OUT OF 84 TESTS
S x T INDEPENDENT OF E =  10. FAILURES OUT OF 78 TESTS


                   OPTIONS - PROTOTYPES
          F1   - GOTO NEXT PROTOYPE
          F2   - CONTINUE RUN
          F3   - PRINT SUMMARY OF AXIOM VIOLATIONS
          F4   - PRINT COMPLETE AXIOM HISTORY
          ESC  - MAIN MENU
```

INDIVIDUAL AXIOMS

SUBJECTS

1 2 3 4 5 6 7 8 9 10 11 12

PRESS THE NUMBER OF THE SUBJECT YOU WANT TO INCLUDE THEN PRESS <RETURN>
PRESS THE 'Q' OR 'q' KEYS FOLLOWED BY A <RETURN> TO QUIT

SUBJECT #

```
*******  SUMMARY OF AXIOM VIOLATIONS  *******

              INDIVIDUAL ANALYSIS
                 SUBJECT #  3

      INDEPENDENCE
T INDEPENDENT OF E AND S =   0. FAILURES OUT OF   0 TESTS
E INDEPENDENT OF T AND S =   0. FAILURES OUT OF   0 TESTS
S INDEPENDENT OF T AND E =   0. FAILURES OUT OF   0 TESTS
      DOUBLE CANCELATION
DOUBLE CANCELLATION IN T x E =   0. FAILURES OUT OF   0 TESTS
DOUBLE CANCELLATION IN E x S =   0. FAILURES OUT OF   0 TESTS
DOUBLE CANCELLATION IN S x T =   0. FAILURES OUT OF   0 TESTS
      JOINT INDEPENDENCE
T x E INDEPENDENT OF S =   0. FAILURES OUT OF   0 TESTS
E x S INDEPENDENT OF T =   0. FAILURES OUT OF   0 TESTS
S x T INDEPENDENT OF E =   0. FAILURES OUT OF   0 TESTS


             OPTIONS - INDIVIDUAL
      F1   - GO TO NEXT INDIVIDUAL
      F2   - CONTINUE RUN
      F3   - PRINT SUMMARY OF AXIOM VIOLATIONS
      F4   - PRINT COMPLETE AXIOM HISTORY
      ESC  - MAIN MENU
```

******* SCALING INFORMATION *******

GROUP SCALE

LAST 5 ITERATIONS

ITERATION	THETA	TAU
76	.01083	.97626
77	.01080	.95846
78	.01078	.97626
79	.01075	.95846
80	.01073	.97626

APPROXIMATE RELATIVE IMPORTANCE
OF EACH FACTOR

23.80 % FOR FACTOR T
33.06 % FOR FACTOR E
42.14 % FOR FACTOR S

THE SCALE VALUES FOR THE ITERATIONS BELOW
ARE PRINTED FROM ITERATION NO. 80

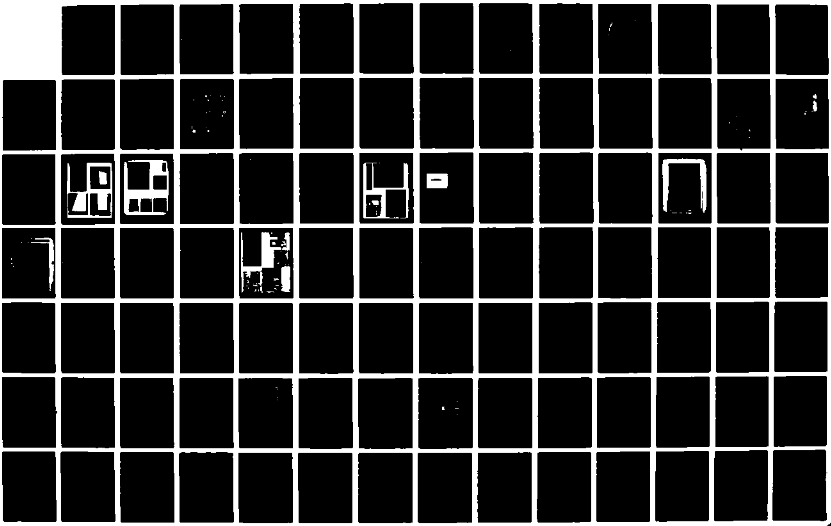
	VARIABLE	ADDITIVE MODEL	ADDITIVE RESCALED
1	TIME 1	-.15	2.78
2	TIME 2	.05	12.35
3	TIME 3	.39	27.58
4	EFFORT 1	-.21	.10
5	EFFORT 2	.15	16.57
6	EFFORT 3	.51	33.16
7	STRESS 1	-.28	-2.88
8	STRESS 2	.15	16.62
9	STRESS 3	.64	39.26

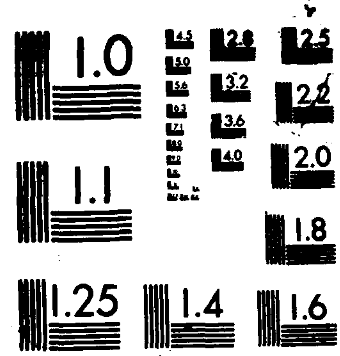
OPTIONS

F1 - PLOT OF RESCALED VS. RAW DATA
F2 - PRINT SCALING INFORMATION
F3 - PRINT ALL ITERATIONS
F4 - VIEW SCALING SOLUTION
F5 - CONTINUE RUN
ESC - MAIN MENU

AD-A185 650 PROCEEDINGS OF THE DOD WORKLOAD ASSESSMENT WORKSHOP ON WORKLOAD ASSESSMEN..(U) NAVAL UNDERWATER SYSTEMS CENTER NEWPORT RI H M FIEDLER 15 SEP 87 NUSC-TD-6608 2/4

UNCLASSIFIED F/G 23/2 NL




1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0

SCALING SOLUTION

F1 - RETURN TO MENU

STIM	LEVELS T	E	S	STANDARD	RESCALED
1	1	1	1	-1.076	.000
2	1	1	2	-.655	17.200
3	1	1	3	-.339	30.100
4	1	2	1	-.774	12.300
5	1	2	2	-.353	29.500
6	1	2	3	-.037	42.400
7	1	3	1	-.333	30.300
8	1	3	2	.087	47.500
9	1	3	3	.403	60.400
10	2	1	1	-.641	17.800
11	2	1	2	-.220	34.900
12	2	1	3	.096	47.800
13	2	2	1	-.339	30.100
14	2	2	2	.082	47.300
15	2	2	3	.398	60.200
16	2	3	1	.102	48.100
17	2	3	2	.523	65.300
18	2	3	3	.839	78.200
19	3	1	1	-.106	39.600
20	3	1	2	.315	56.800

BACKGROUND, DESCRIPTION, AND APPLICATION OF THE NASA TASK LOAD INDEX

NASA-TLX

SANDRA G. HART
NASA-AMES RESEARCH CENTER
MOFFETT FIELD, CA

ASSUMPTION 1:

WORKLOAD IS AN IMPORTANT, PRACTICALLY RELEVANT, MEASURABLE ENTITY. SUBJECTIVE RATINGS ARE THE MOST COMMONLY-USED ASSESSMENT PROCEDURE, PROVIDE THE MOST DIRECT INFORMATION ABOUT THE IMPACT OF TASKS ON AN OPERATOR, AND INTEGRATE THE EFFECTS OF MANY WORKLOAD CONTRIBUTORS

IMPLICATION:

DEVELOPING A SENSITIVE AND RELIABLE SUBJECTIVE WORKLOAD ASSESSMENT PROCEDURE WAS THE FIRST PRIORITY OF THE NASA WORKLOAD ASSESSMENT PROGRAM

SUBJECTIVE WORKLOAD ASSESSMENT PROCEDURES: PROBLEMS

- HIGH BETWEEN-SUBJECT VARIABILITY
- WORKLOAD SOURCES VARY ACROSS TASKS
- LIMITED TO REMEMBERED EVENTS, ACTIVITIES, FEELINGS
- ABSOLUTE VALUES ARE NOT MEANINGFUL ACROSS TASKS
- RATER BIASES, DIFFERENCES IN DEFINITION
- EXTERNAL VALIDATION DIFFICULT
  - WORKLOAD NOT UNIQUELY DEFINED BY TASK DEMANDS
  - NO OBJECTIVE STANDARDS AGAINST WHICH SCALES CAN BE EVALUATED
- SCALING PROBLEMS (ANCHORS, INCREMENTS, FORMAT)

SUBJECTIVE WORKLOAD EVALUATION DEVELOPMENT EFFORT: GOALS

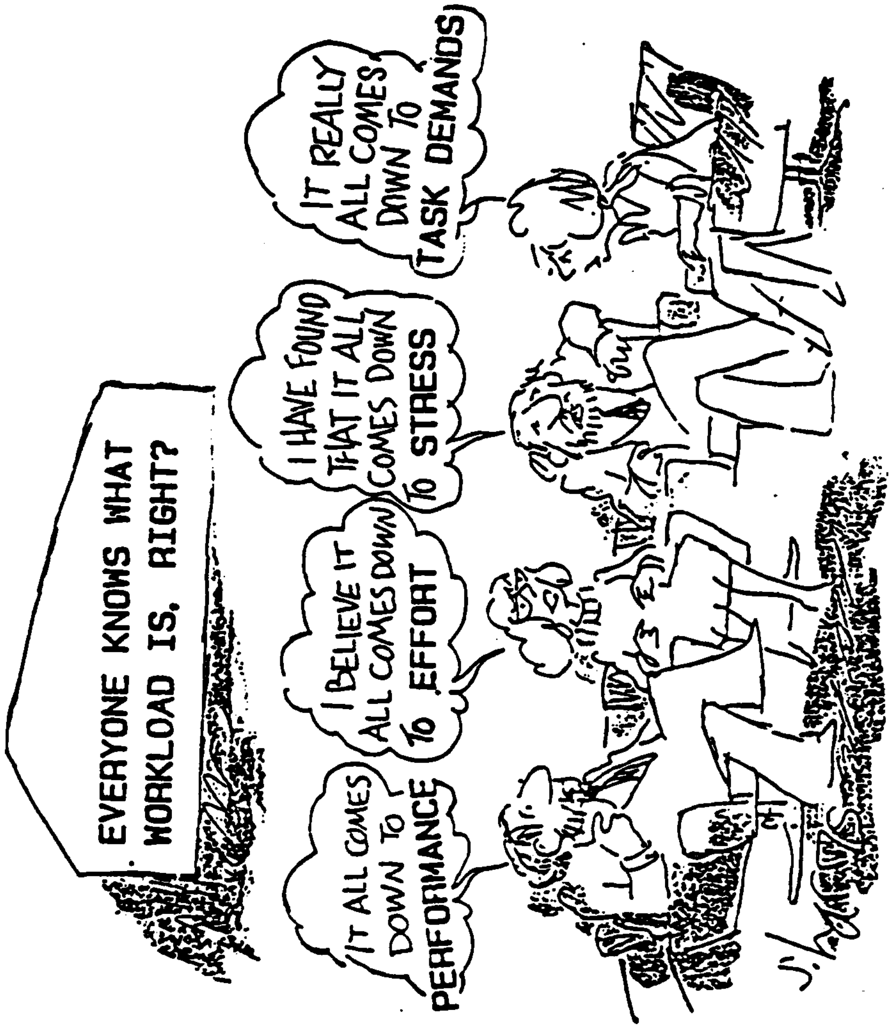
- IDENTIFY THE FACTORS THAT INFLUENCE THE SUBJECTIVE EXPERIENCE OF WORKLOAD
- DEFINE A MINIMUM SET OF SUBSCALES
- DEVELOP A METHOD OF COMPUTING AN ESTIMATE OF OVERALL WORKLOAD FROM THE SUBSCALE RATINGS
- EVALUATE THE RATING TECHNIQUE

ASSUMPTION 2:

ALTHOUGH EVERYONE HAS AN INTUITIVE IDEA ABOUT WHAT WORKLOAD IS, THE FACTORS EACH ONE MIGHT INCLUDE IN A PERSONAL EXPERIENCE OR A FORMAL DEFINITION VARY

IMPLICATION:

SUBJECTIVE RATING SCALES MUST ADJUST FOR PERSONAL BIASES


EVERYONE KNOWS WHAT WORKLOAD IS, RIGHT?
IT ALL COMES DOWN TO PERFORMANCE
I BELIEVE IT ALL COMES DOWN TO EFFORT
I HAVE FOUND THAT IT ALL COMES DOWN TO STRESS
IT REALLY ALL COMES DOWN TO TASK DEMANDS

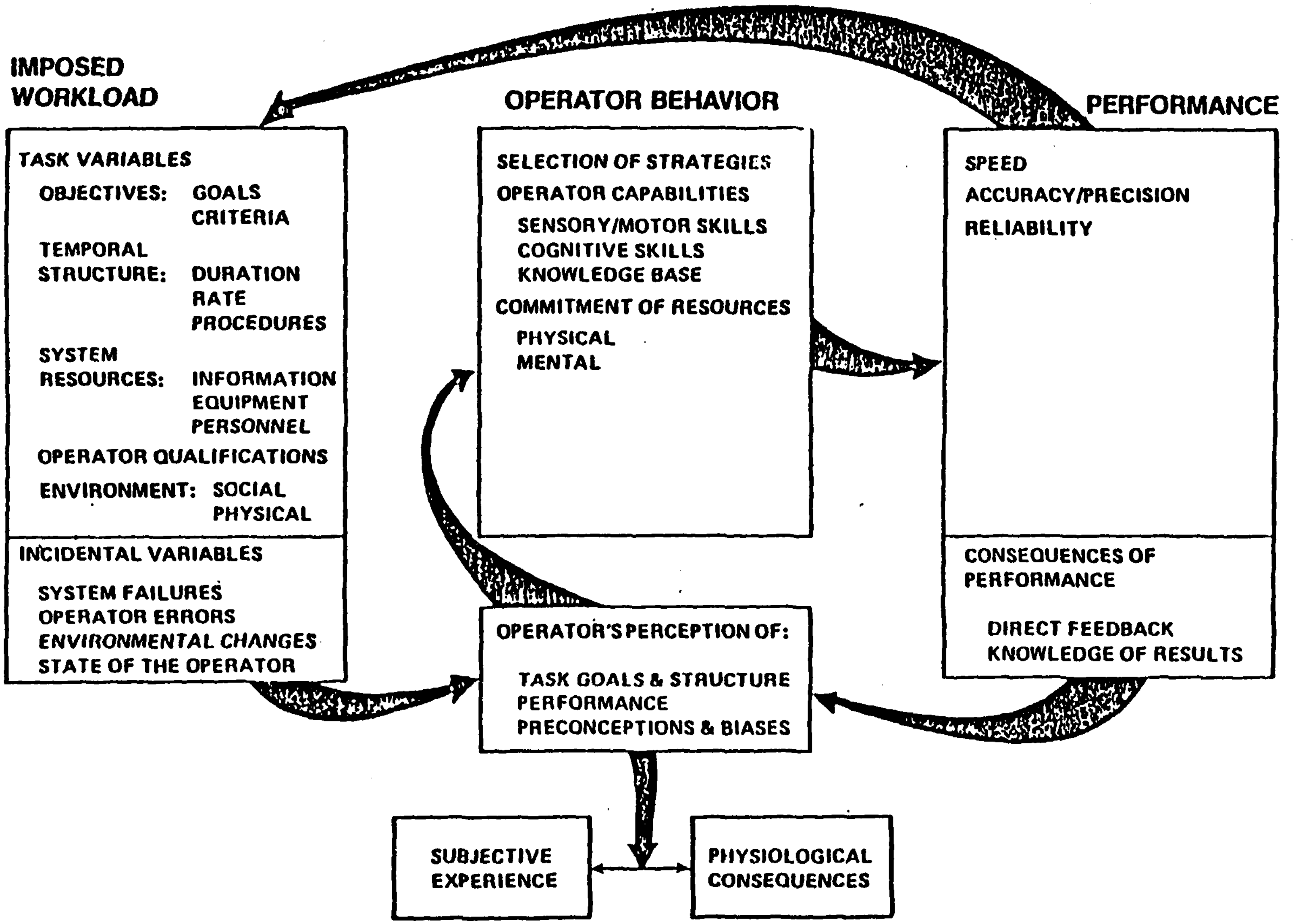
ASSUMPTION 3:

WORKLOAD IS A MULTI-DIMENSIONAL CONSTRUCT THAT REFLECTS THE INTERSECTION BETWEEN A PARTICULAR OPERATOR PERFORMING A SPECIFIC SET OF TASK DEMANDS IN A GIVEN ENVIRONMENT

IMPLICATION:

SUBJECTIVE RATING SCALES MUST ENCOMPASS TASK-, OPERATOR-, ENVIRONMENT-, AND PERFORMANCE-RELATED FACTORS

CONCEPTUAL FRAMEWORK FOR THE ANALYSIS OF WORKLOAD AND PERFORMANCE



ASSUMPTION 4:

RATERS CAN QUANTIFY THE MAGNITUDES OF COMPONENT FACTORS MORE CONSISTENTLY AND ACCURATELY THAN THEY CAN THE MORE GLOBAL CONCEPT OF WORKLOAD

IMPLICATION:

A SUBJECTIVE RATING SCALE MUST INCLUDE MULTIPLE SUBSCALES

ASSUMPTION 5:

TRANSLATING A PERSONAL EXPERIENCE OF WORKLOAD INTO A FORMAL EVALUATION IS NOT A COMMON OR NATURAL REQUIREMENT

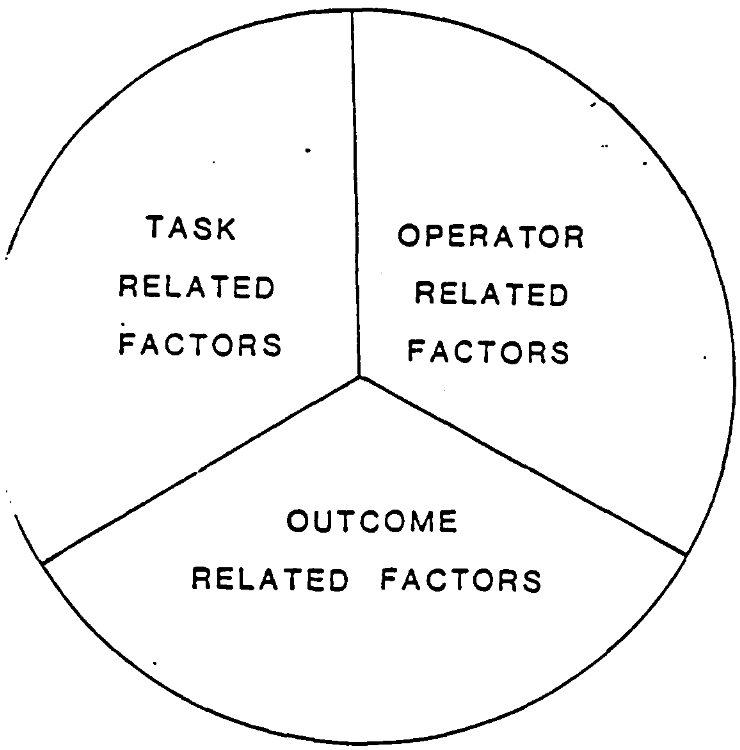
IMPLICATION:

SUBJECTIVE RATING TECHNIQUES MUST ALLOW RATERS TO EXPRESS THEIR EVALUATIONS AS NATURALLY AS POSSIBLE (E.G. VERBAL DESCRIPTIONS, MAGNITUDE JUDGEMENTS)

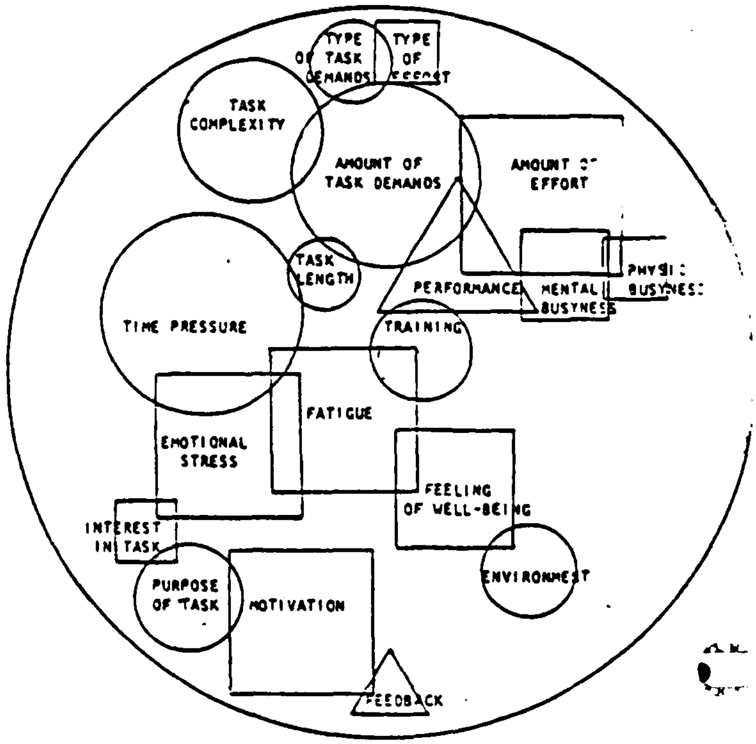
PHASE I

- SURVEY AND ANALYZE INDIVIDUAL DEFINITIONS OF WORKLOAD
- PERFORM INITIAL SELECTION OF SUBSCALES

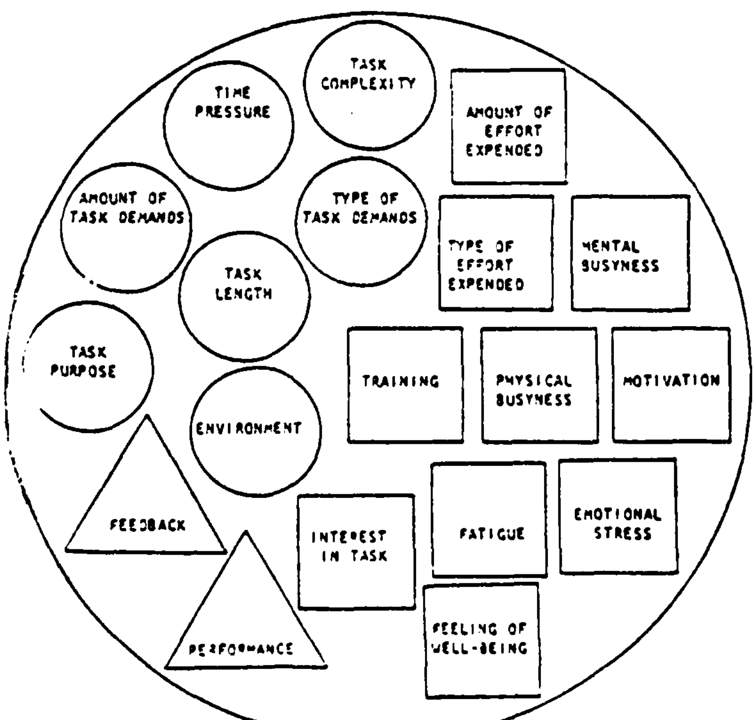
WORKLOAD



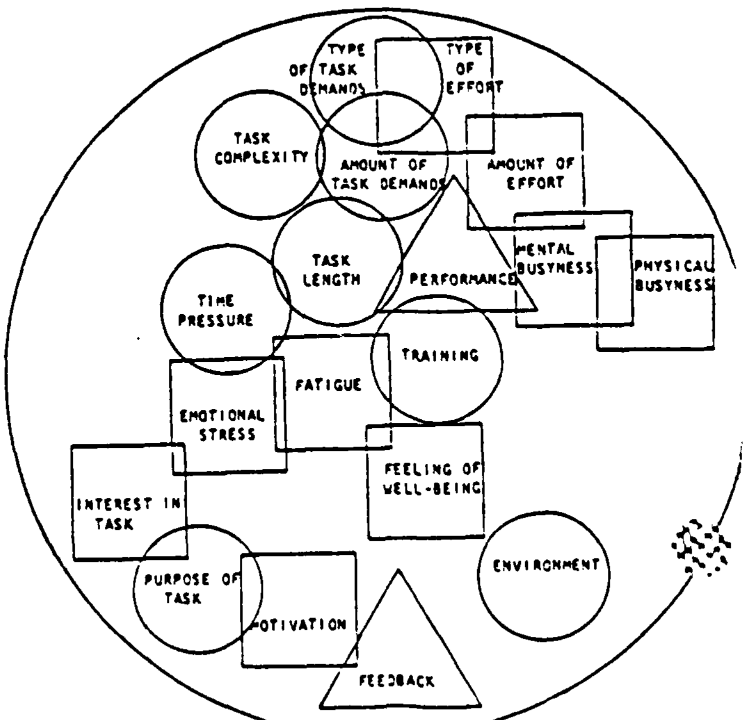
THE IMPORTANCE PLACED ON EACH FACTOR WHEN ESTIMATING OR REACTING TO WORKLOAD VARIES FROM ONE INDIVIDUAL TO ANOTHER.



MULTIPLE FACTORS WITHIN EACH DIMENSION MAY DIRECTLY OR INDIRECTLY CONTRIBUTE TO THE EXPERIENCE OF WORKLOAD



THE FACTORS ARE NOT NECESSARILY RELEVANT IN EVERY SITUATION, AND THEY MAY OR MAY NOT OPERATE INDEPENDENTLY.



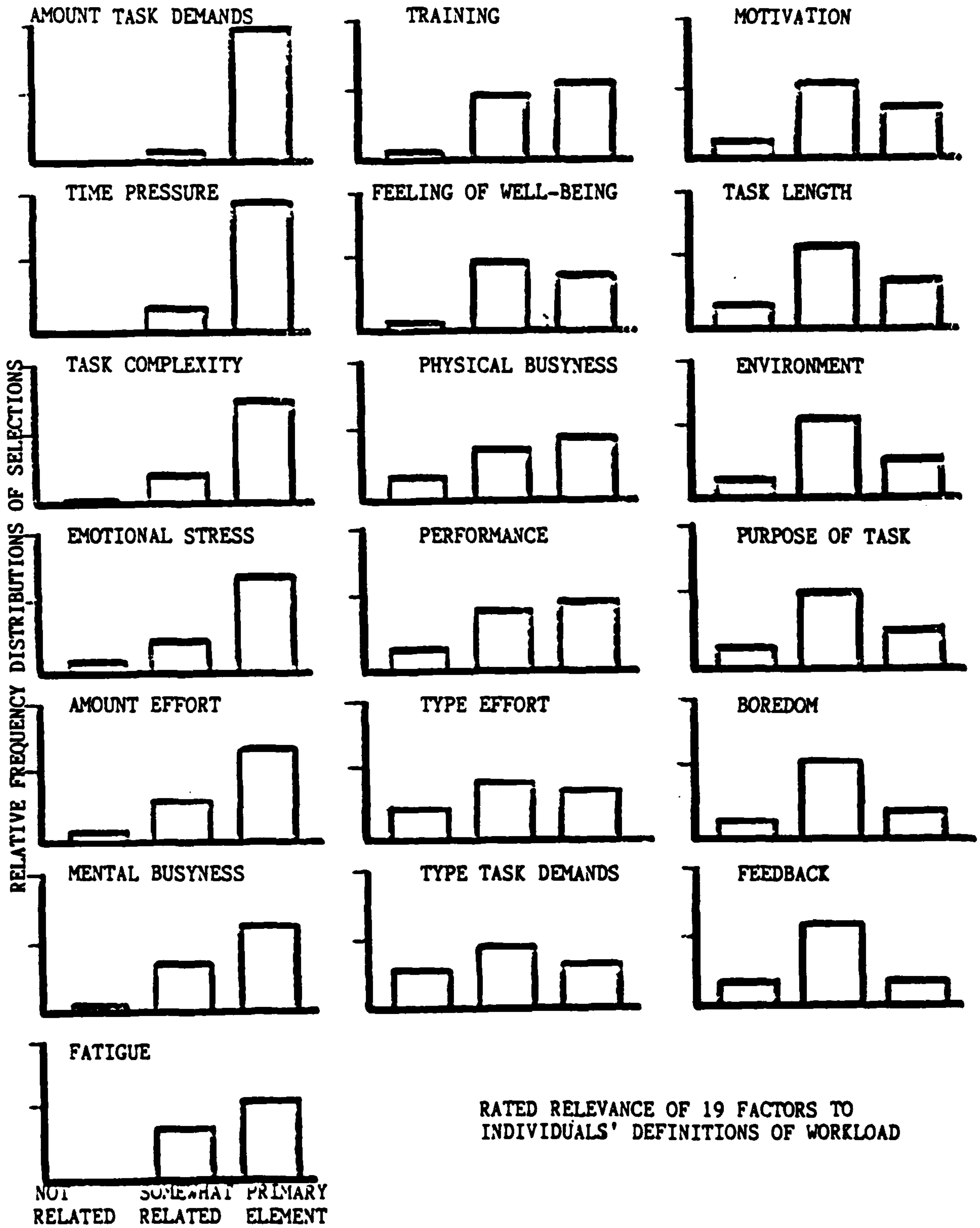
METHOD OF EVALUATING THE FACTORS THAT ARE INCLUDED IN INDIVIDUALS' WORKLOAD DEFINITIONS

FACTORS:

TASK DEMANDS-TYPE	PERFORMANCE	FEELING OF WELL-BEING
TASK DEMANDS-AMOUNT	FEEDBACK GIVEN	FATIGUE
TASK-PURPOSE	ENVIRONMENT	EMOTIONAL STRESS
TASK-COMPLEXITY	TRAINING LEVEL	BOREDOM
TASK-LENGTH	MENTAL BUSYNESS	MOTIVATION
TASK-TIME PRESSURE	PHYSICAL BUSYNESS	EFFORT-AMOUNT
	EFFORT-TYPE	

DECISIONS:

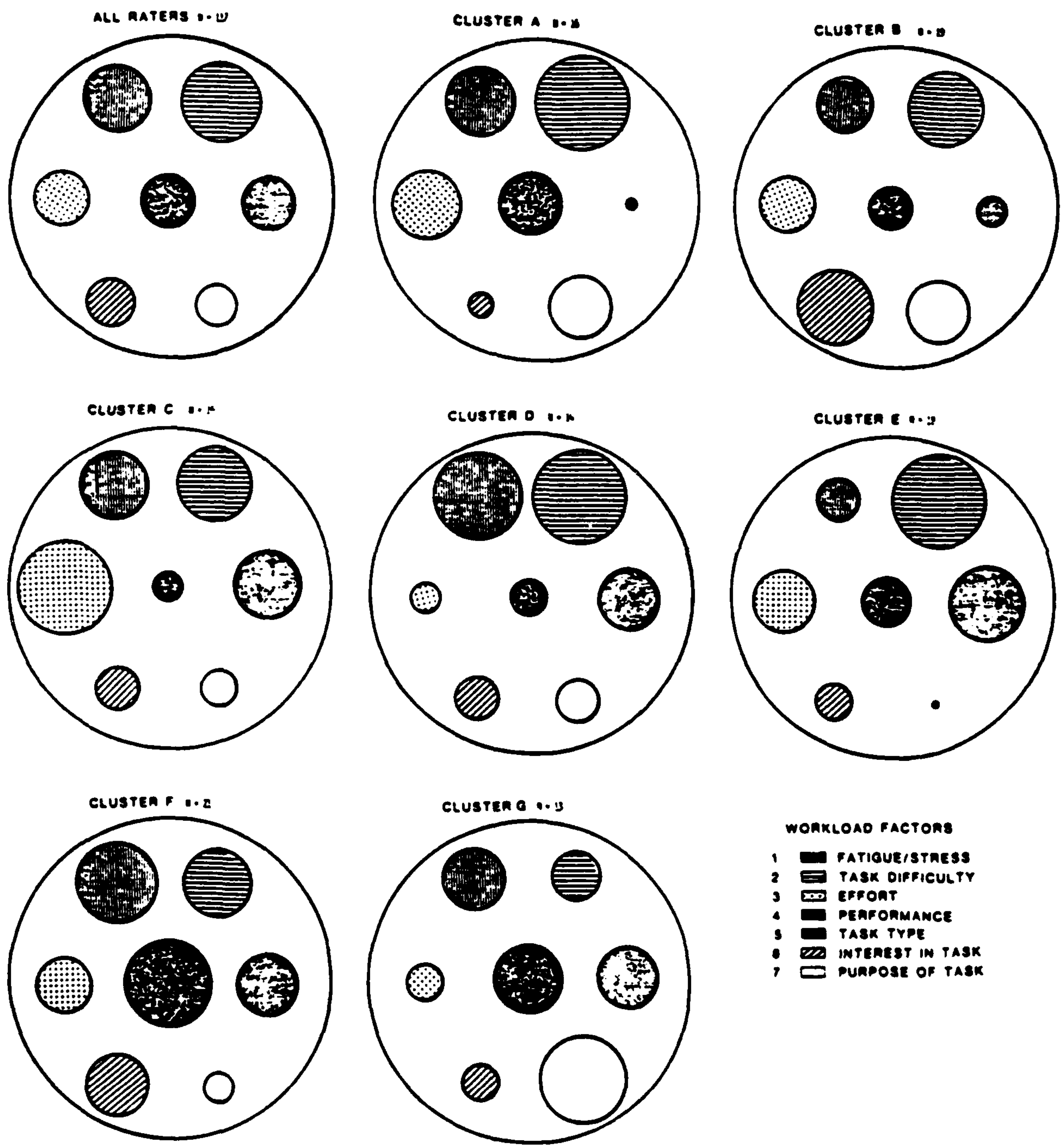
THE FACTOR IS NOT AT ALL RELATED TO WORKLOAD	THE FACTOR IS SOME-WHAT RELATED TO WORKLOAD	THE FACTOR IS A PRIMARY COMPONENT OF WORKLOAD



RATED RELEVANCE OF 19 FACTORS TO INDIVIDUALS' DEFINITIONS OF WORKLOAD

PATTERNS OF RATINGS FOR THE 7 FACTORS IDENTIFIED AS WORKLOAD COMPONENTS

(THE DIAMETER OF EACH CIRCLE IS PROPORTIONAL TO RATINGS GIVEN TO COMPONENTS WITH FACTOR LOADINGS GREATER THAN .5)



PRELIMINARY SET OF SUBSCALES

TASK DIFFICULTY	ATTENTION	STRESS
TASK COMPLEXITY	MENTAL EFFORT	ENERGY LEVEL
TIME PRESSURE	SENSORY EFFORT	PHYSICAL STATE
ACTIVITY LEVEL	PHYSICAL EFFORT	MOTIVATION
TRAINING	PERFORMANCE	

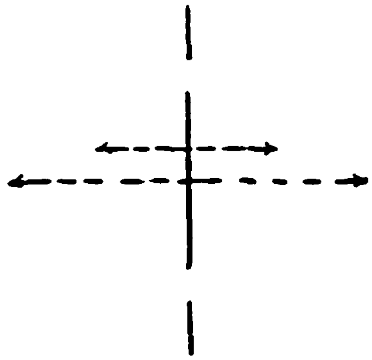
PHASE II

- PERFORM INITIAL EXPERIMENTAL EVALUATION OF 15 SUBSCALES TO DETERMINE
    - SENSITIVITY TO EXPERIMENTALLY MANIPULATED WORKLOAD LEVELS
    - RELATIONSHIP WITH OVERALL WORKLOAD RATINGS
    - INDEPENDENCE/ASSOCIATION AMONG SUBSCALES
- SELECT NINE CANDIDATE SUBSCALES

WORKLOAD TASKS

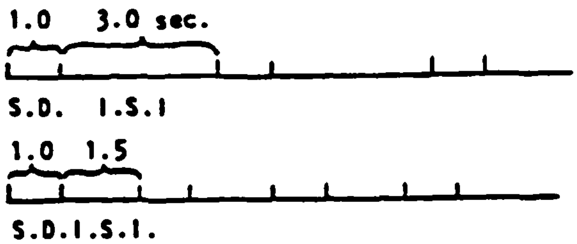
TRACKING TASK

SINGLE AXIS
COMPENSATORY
K/S PLANT



	BAND WIDTH 1.0	BAND WIDTH 2.0
STANDARD DEVIATION 32		
STANDARD DEVIATION 64		

STERNBERG MEMORY TASK



SLOW VERSION 20 CHARACTERS/MINUTE
FAST VERSION 40 CHARACTERS/MINUTE

	SPEED: SLOW	SPEED: FAST
MEMORY LOAD 1		
MEMORY LOAD 5		

AUDITORY MONITORING TASK

SLOW VERSION 8 LOW, 6 MEDIUM, 4 HIGH PER MINUTE

FAST VERSION 16 LOW, 12 MEDIUM, 8 HIGH PER MINUTE

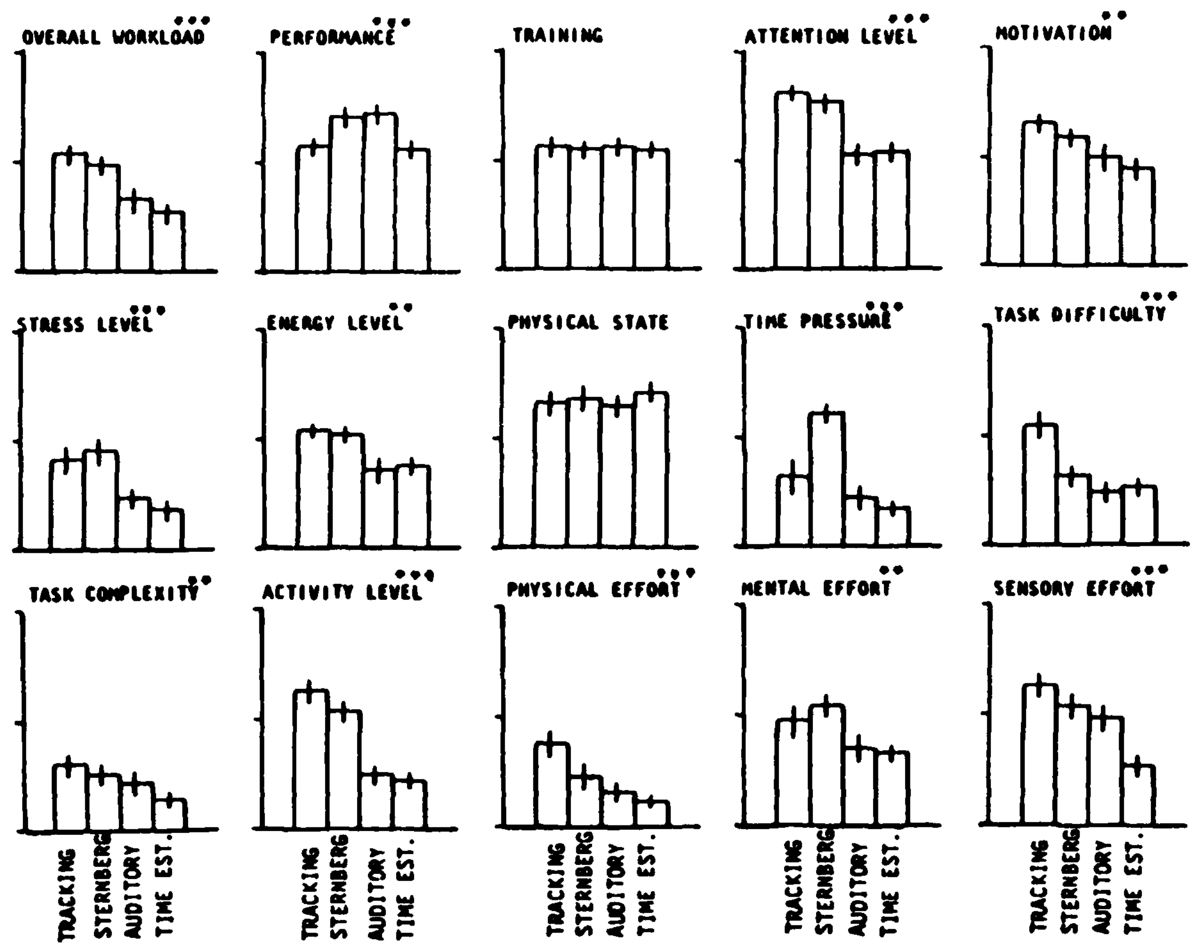
	SPEED: SLOW	SPEED: FAST
NO MEMORY LOAD	EVERY LOW 8 RESPONSES	EVERY HIGH 8 RESPONSES
MEMORY LOAD	EVERY 2ND HIGH 2 RESPONSES	EVERY 4TH HIGH 2 RESPONSES

TIME ESTIMATION TASK

PRODUCTION	ESTIMATION
BEGIN PRODUCTION n SECONDS	BEGIN INTERVAL
PRODUCTION n SECONDS	INTERVAL
END OF PRODUCTION n SECONDS	END INTERVAL
	ESTIMATE

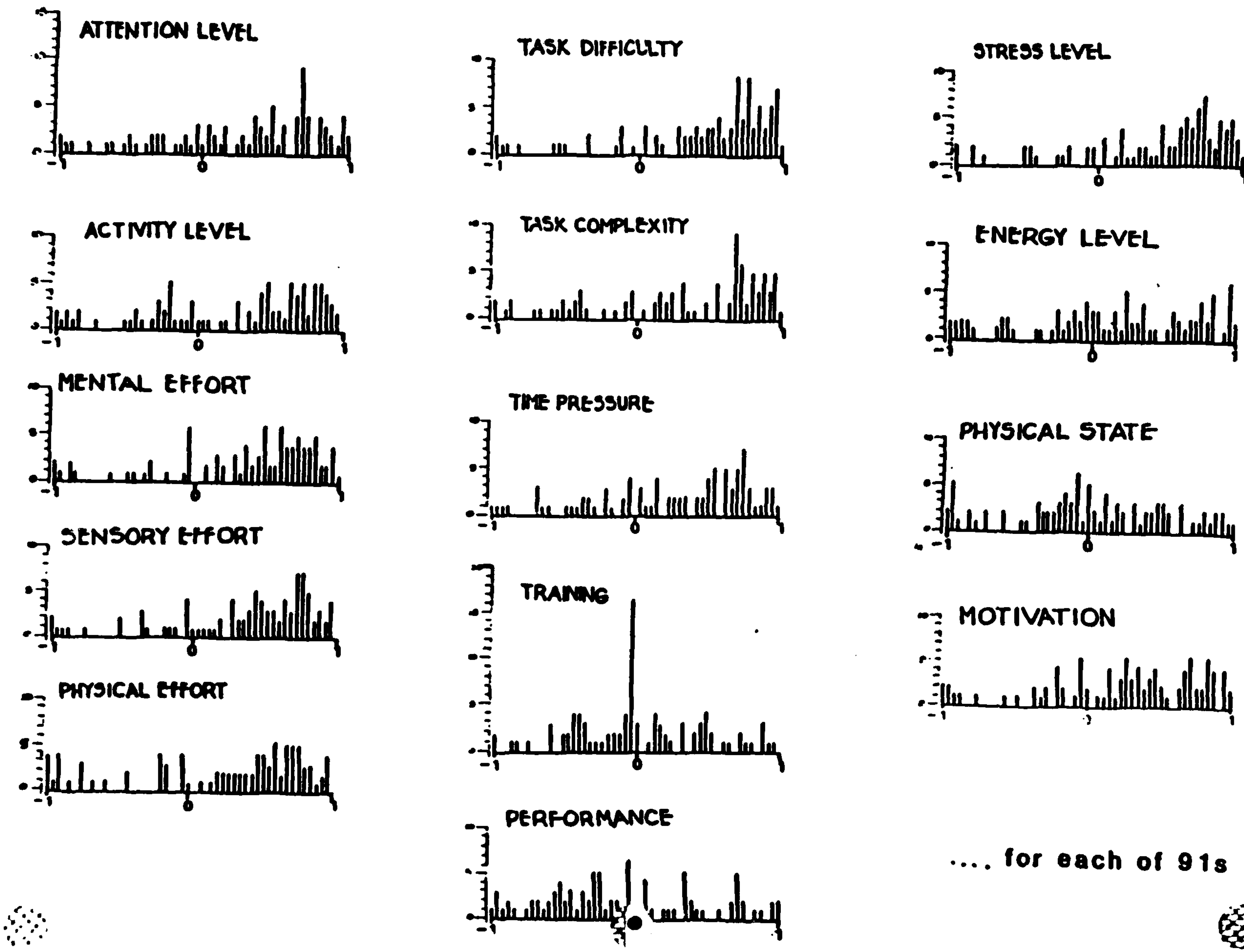
INTERVALS OF 5-14 SECONDS
20 INTERVALS/SESSION

	TECHNIQUE: NO COUNT	TECHNIQUE: COUNT
PRODUCTION METHOD		
VERBAL ESTIMATION METHOD		



Rating scale means and standard errors calc ıted over the four levels of each task with significant between task differences .dentified ** p<.01, *** p<.001.

FREQUENCY DISTRIBUTION OF Rxy COEFFICIENTS OVERALL WORKLOAD vs:



NASA BIPOLAR RATING SCALE: SUBSCALES

TASK DIFFICULTY

TIME PRESSURE

OWN PERFORMANCE

PHYSICAL EFFORT

MENTAL EFFORT

EMOTIONAL STRESS

FRUSTRATION

FATIGUE

ACTIVITY TYPE

NOTE: OVERALL WORKLOAD RATINGS WERE OBTAINED AS WELL FOR SCALE EVALUATION

PHASE III

- EXPERIMENTALLY EVALUATE THE NINE SUBSCALES AND THE PROPOSED WEIGHTING AND AVERAGING PROCEDURE (THE NASA BIPOLAR RATING SCALE)
  - SENSITIVITY TO EXPERIMENTAL'Y MANIPULATED WORKLOAD LEVELS
  - RELATIONSHIP WITH OVERALL WORKLOAD RATINGS
  - INDEPENDENCE/ASSOCIATION AMONG SUBSCALES
  - RELATIONSHIP BETWEEN WEIGHTS AND RATINGS
- COMPARE THIS TECHNIQUE WITH OTHER METHODS:
  - SENSITIVITY
  - BETWEEN-SUBJECT VARIABILITY
  - EASE OF USE

SUBJECTIVE RATINGS

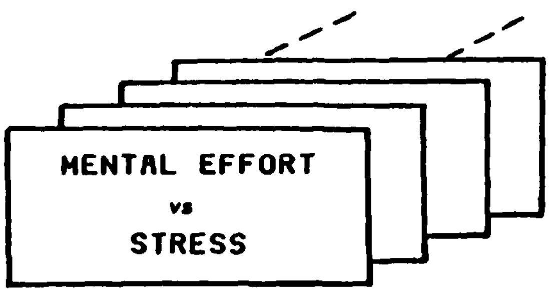
METHOD FOR REDUCING THE BETWEEN-SUBJECT VARIABILITY

WORKLOAD DIMENSIONS:

- TASK DIFFICULTY
- TIME PRESSURE
- OWN PERFORMANCE
- PHYSICAL EFFORT
- MENTAL EFFORT
- STRESS
- FATIGUE
- FRUSTRATION
- TYPE OF TASK

1. "WEIGHTS"

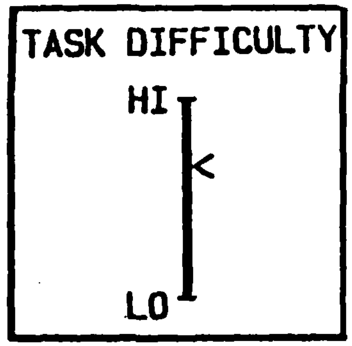
EACH OF 9 FACTORS IS COMPARED WITH EVERY OTHER ONE (WHICH IS MORE RE-LATED TO WORKLOAD?)



0 = NEVER SELECTED
8 = ALWAYS SELECTED

2. BIPOLAR RATINGS:

THE AMOUNT OF EACH FACTOR EXPERIENCED IN A TASK IS EVALUATED ON A BIPOLAR SCALE:

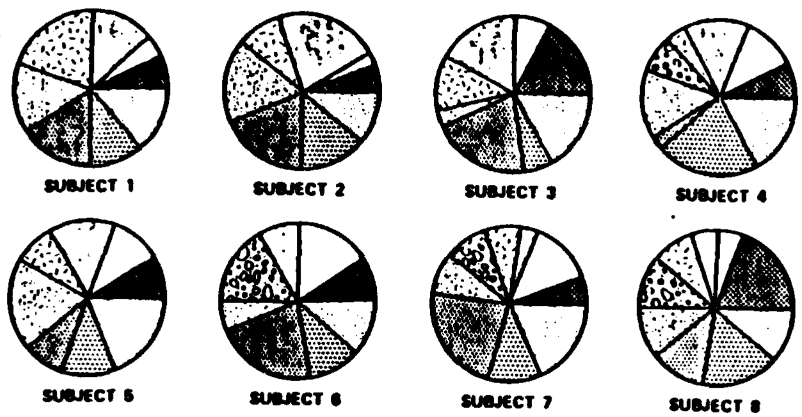


3. WEIGHTING PROCEDURE

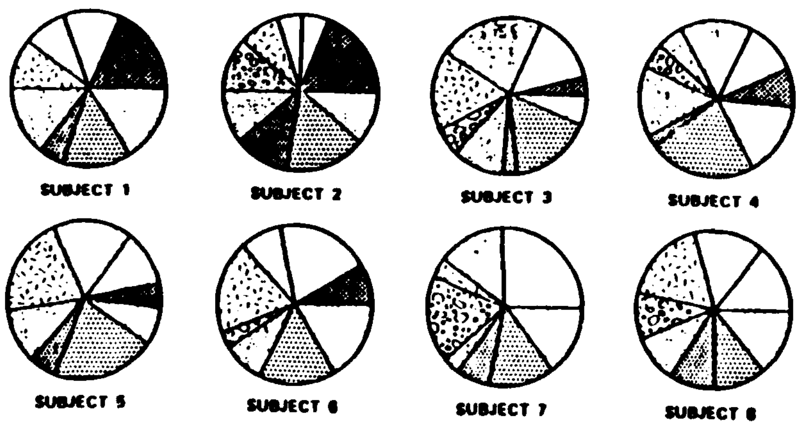
EACH "RATING" IS WEIGHTED BY ITS SUBJECTIVE IMPORTANCE TO EACH SUBJECT (THE "WEIGHTS")

THE AVERAGE OF THE WEIGHTED RATINGS = DERIVED WORKLOAD SCORE

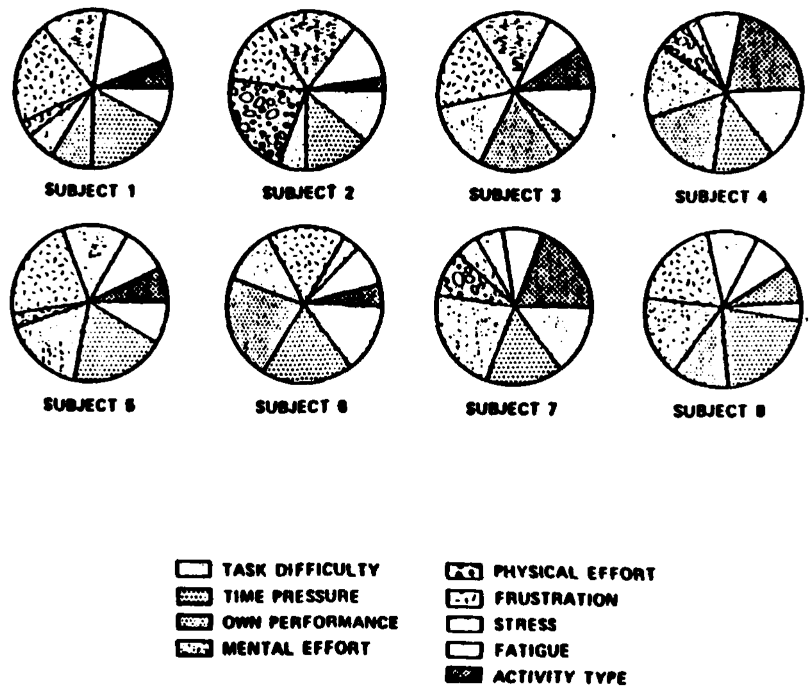
THE RELATIVE IMPORTANCE TO EXPERIENCED WORKLOAD OF 9 FACTORS AS REPORTED BY EACH OF 8 PILOTS
(36 PAIRWISE COMPARISONS)



THE RELATIVE IMPORTANCE TO EXPERIENCED WORKLOAD OF 9 FACTORS AS REPORTED BY EACH OF 8 PILOTS
(36 PAIRWISE COMPARISONS)



THE RELATIVE IMPORTANCE TO EXPERIENCED WORKLOAD OF 9 FACTORS AS REPORTED BY EACH OF 8 PILOTS
(36 PAIRWISE COMPARISONS)

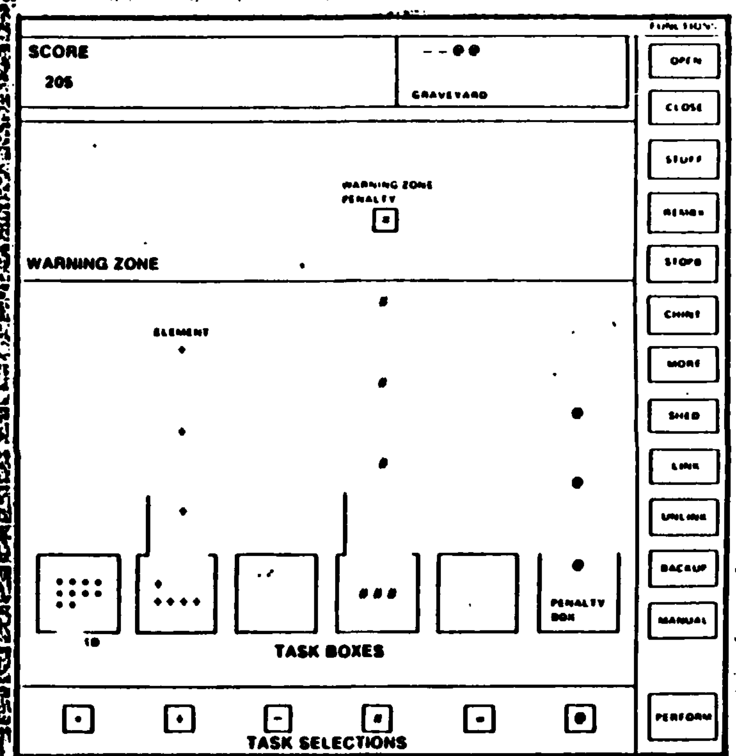


CRITERION TASK: POPCORN

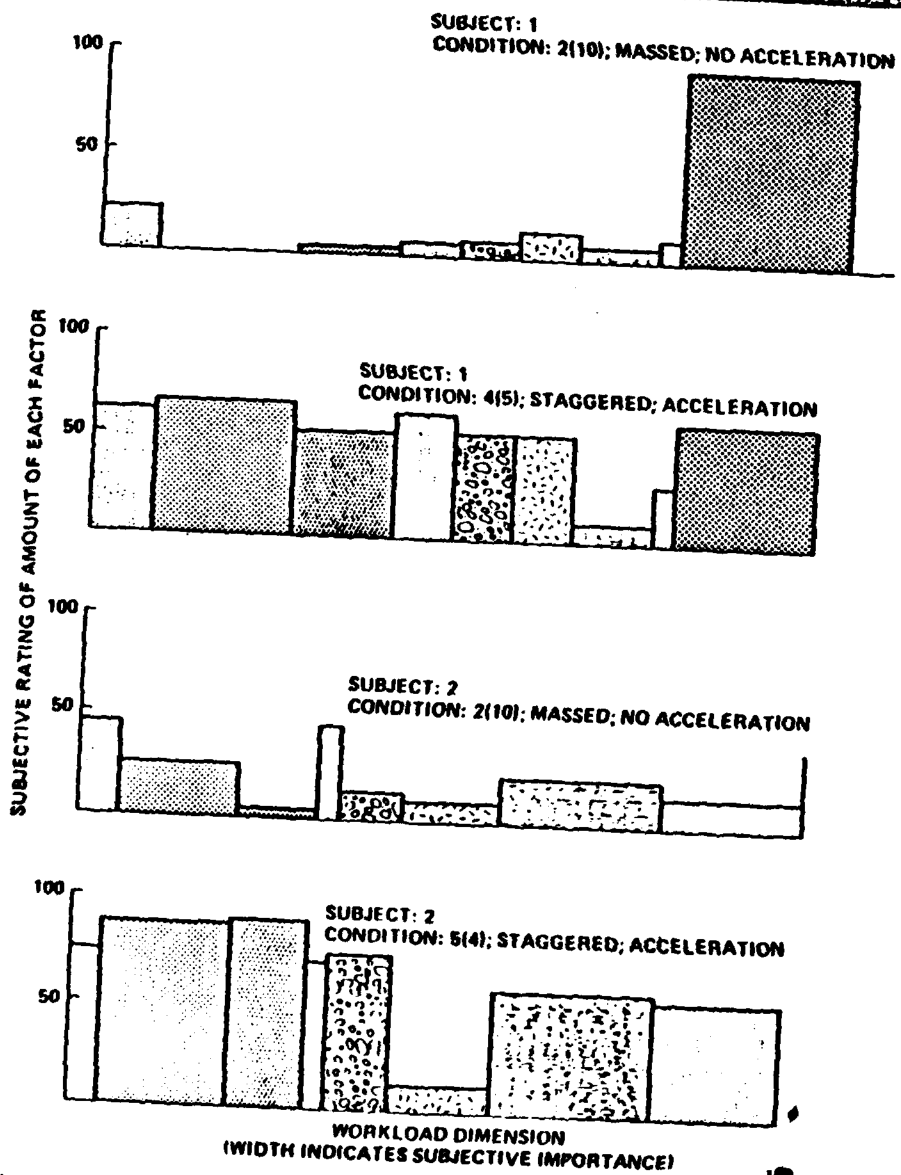
OBJECTIVE: SIMULATE SUPERVISORY CONTROL ENVIRONMENT IN WHICH MULTIPLE, CONCURRENT TASKS ARE ACCOMPLISHED WITH AUTOMATIC SUBSYSTEMS

APPLICATIONS:

- o STUDY TIME-PRESSURE ASPECTS OF WORKLOAD
- o COMPARE WORKLOAD PREDICTIONS TO ASSESSMENTS
- o PROVIDE DATA BASE FOR MODELS OF HUMAN PERFORMANCE
- o EXAMINE INDIVIDUAL DIFFERENCES (E.G., TYPE A vs TYPE B)
- o COGNITIVE TEST USED BY USAF IN RETURNING PILOTS TO FLIGHT STATUS



SUBJECTIVE RATINGS: GRAPHIC EXAMPLE OF WEIGHTING PROCEDURE



EACH BAR REPRESENTS ONE WORKLOAD-RELATED FACTOR

BAR WIDTH = IMPORTANCE OF FACTOR TO THE RATER

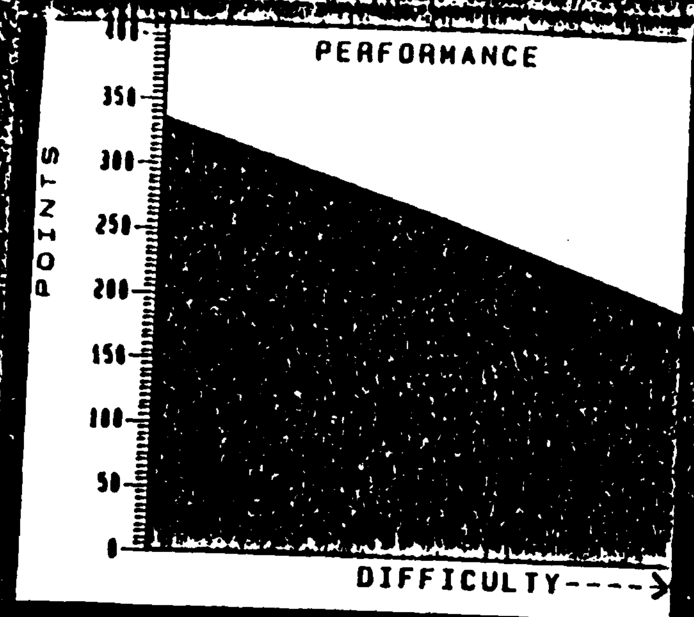
BAR HEIGHT = SUBJECTIVE MAGNITUDE OF THE FACTOR IN A TASK

THE AREA OF EACH GRAPH = WORKLOAD EXPERIENCE OF RATER

WORKLOAD-RELATED FACTORS:

- TASK DIFFICULTY
- TIME PRESSURE
- OWN PERFORMANCE
- MENTAL EFFORT
- PHYSICAL EFFORT
- FRUSTRATION
- STRESS
- FATIGUE
- ACTIVITY TYPE

Hart & Battiste, 1984

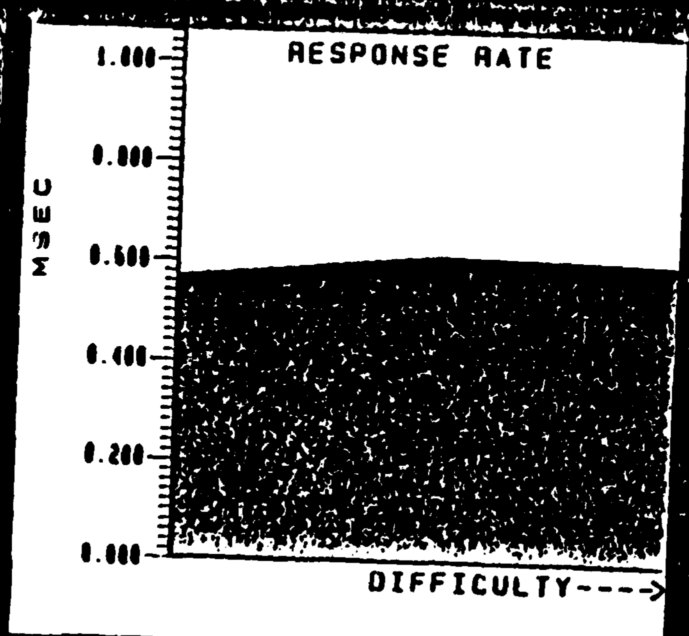


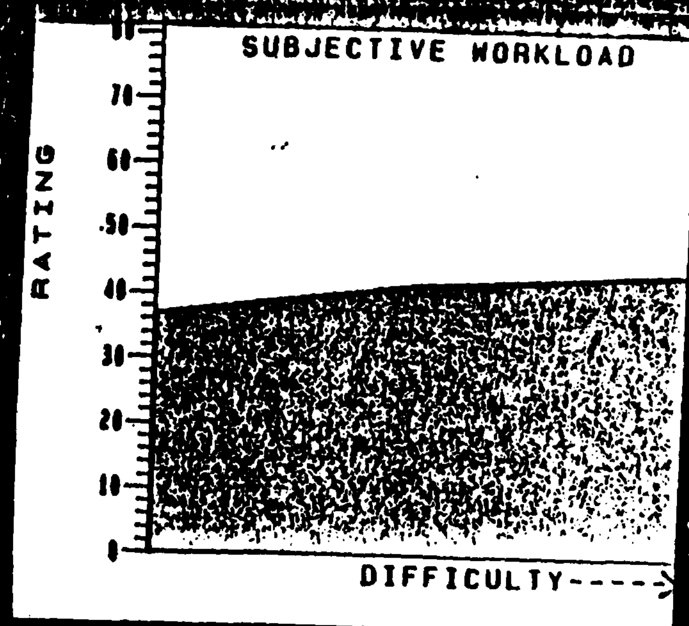
MEASURES OF PERFORMANCE AND WORKLOAD MAY NOT COVARY

WORKLOAD MEASURES MAY REFLECT THE EFFORT EXERTED TO ACCOMPLISH TASK REQUIREMENTS

PERFORMANCE MEASURES REFLECT THE ADEQUACY OF THE EFFORT RELATIVE TO AN OBJECTIVE CRITERIA.

EXAMPLE: MEASURES OF PERFORMANCE (SCORE), BEHAVIOR (RESPONSE RATE) AND WORKLOAD (RATINGS) OBTAINED FOR THREE LEVELS OF DIFFICULTY IMPOSED BY A SUPERVISORY CONTROL SIMULATION.





OBJECTIVE

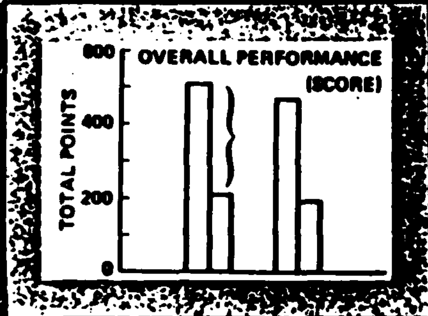
PROVIDE EMPIRICAL VALIDATION OF HYPOTHESIS THAT "TYPE A" INDIVIDUALS ARE MORE REACTIVE TO TASK-INDUCED STRESSORS

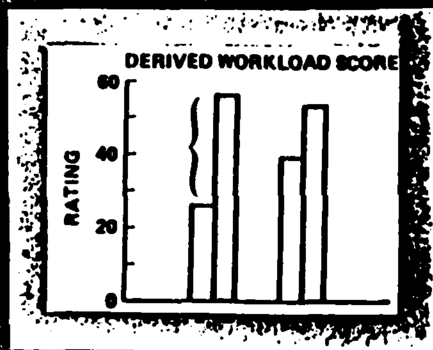
APPROACH

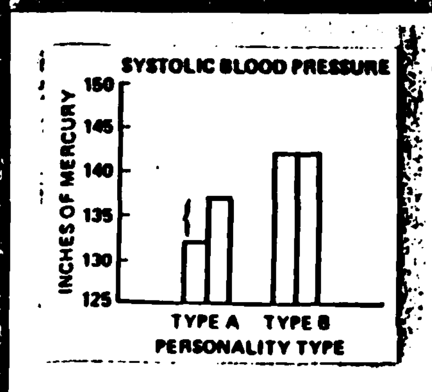
MEASURE HEART RATE, BLOOD PRESSURE SUBJECTIVE RATINGS STRATEGY SHIFTS' AND PERFORMANCE IN RESPONSE TO EXPERIMENTAL MANIPULATIONS

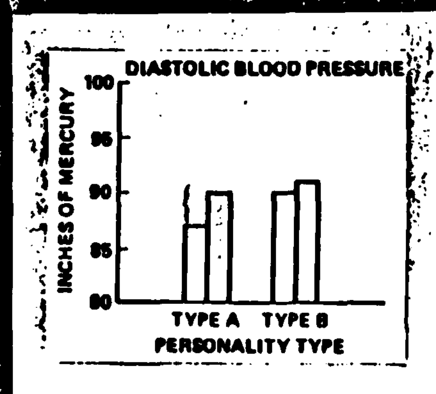
RESULTS

TYPE A MEN ARE SIGNIFICANTLY MORE REACTIVE, PHYSIOLOGICALLY, SUBJECTIVELY, AND BEHAVIORALLY, THAN TYPE B MEN









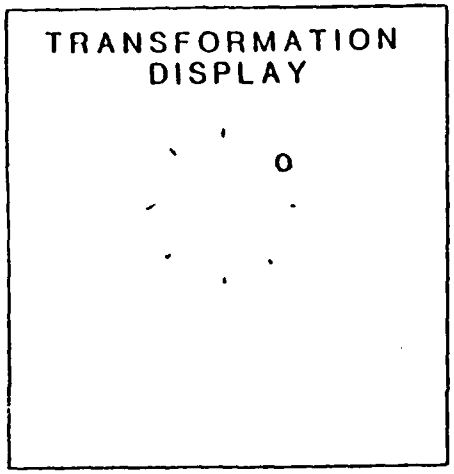
EXPERIMENTAL TASK

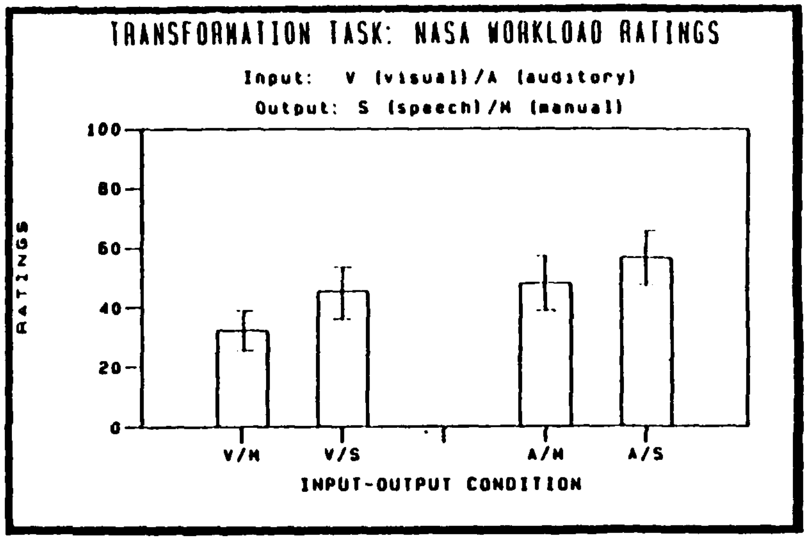
EASIEST SCENARIO

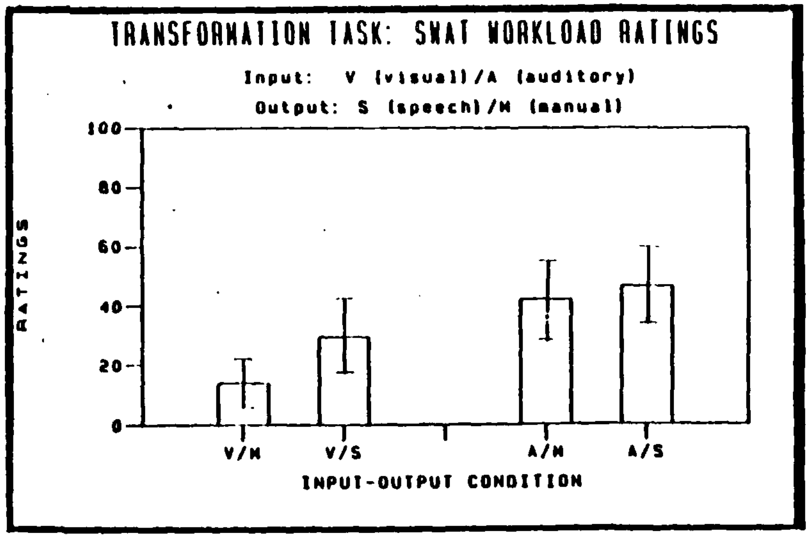
HARDEST SCENARIO

123

COMPARISON OF SUBJECTIVE RATING TECHNIQUES (2)







Vidulich & Tsang, 1985

OBJECTIVE: TO COMPARE THE SENSITIVITY AND STABILITY OF TWO, MULTI-DIMENSIONAL RATING TECHNIQUES FOR WORKLOAD ASSESSMENT

APPROACH: TRANSFORMATION TASKS
RATINGS WERE OBTAINED WITH EACH TECHNIQUE FOLLOWING THE PERFORMANCE OF A SPATIAL TRANSFORMATION TASK:
(E.G. "NORTH-EAST", "SOUTH", etc)

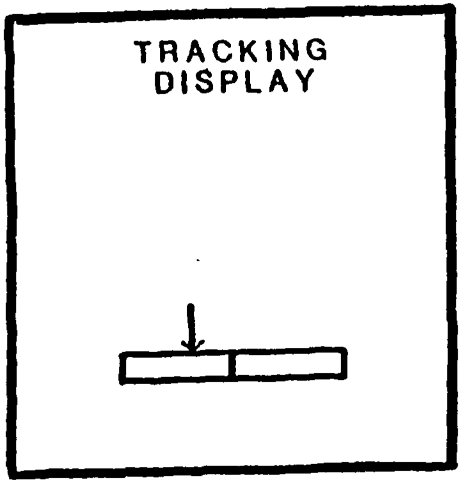
RESULTS:

FOR BOTH TECHNIQUES, AUDITORY INPUT WAS RATED AS SIGNIFICANTLY MORE LOADING THAN VISUAL

FOR BOTH TECHNIQUES, SPEECH OUTPUT WAS RATED AS SIGNIFICANTLY MORE LOADING THAN MANUAL

AGAIN, BETWEEN-SUBJECT VARIABILITY WAS GREATER WITH THE SWAT TECHNIQUE

COMPARISON OF SUBJECTIVE RATING TECHNIQUES (1)

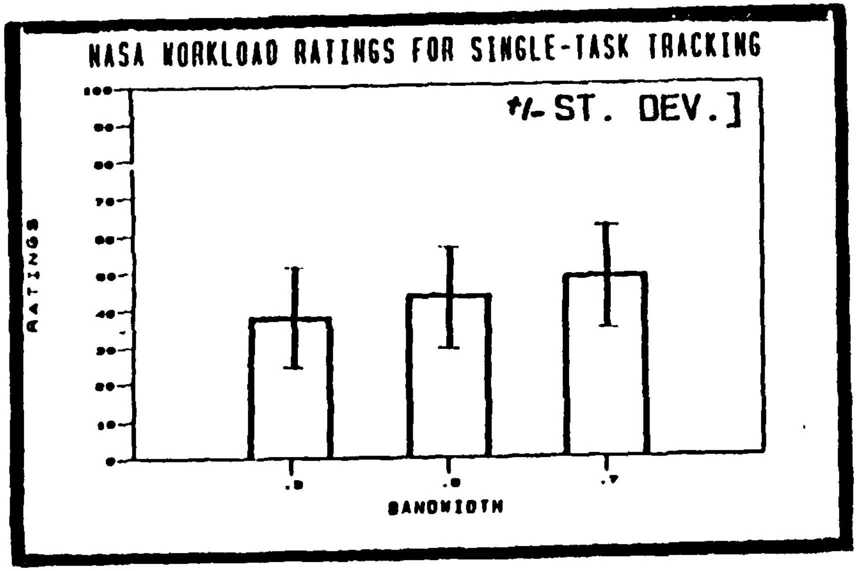


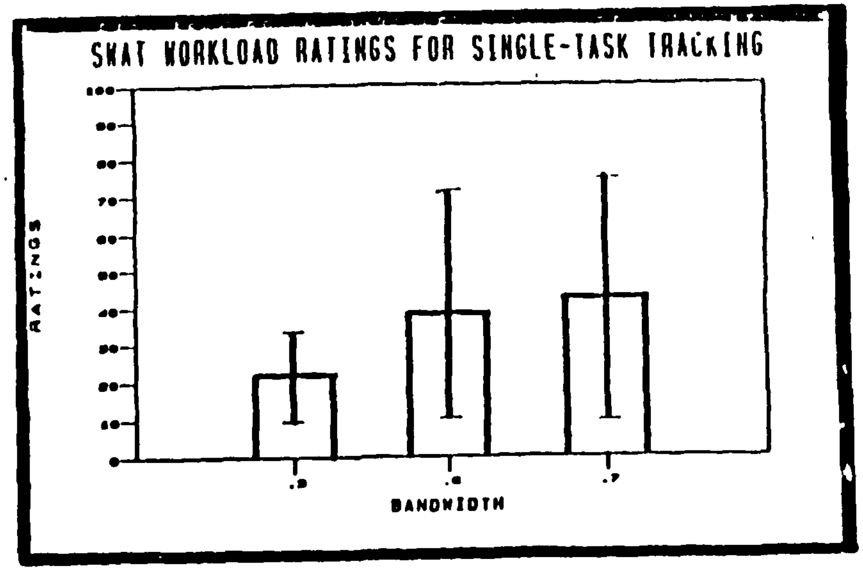
OBJECTIVE: TO COMPARE THE SENSITIVITY AND STABILITY OF TWO MULTI-DIMENSIONAL RATING TECHNIQUES FOR WORKLOAD ASSESSMENT

APPROACH: TRACKING TASKS
RATINGS WERE OBTAINED WITH EACH TECHNIQUE FOLLOWING THE PERFORMANCE OF A SINGLE-AXIS TRACKING TASK:
BANDWIDTHS = .3, .5, .7

RESULTS: ALTHOUGH BOTH MEASURES INDICATED INCREASING WORKLOAD AS BANDWIDTH INCREASED, THE DIFFERENCE WAS STATISTICALLY SIGNIFICANT FOR THE NASA RATINGS ONLY.

BETWEEN-SUBJECT VARIABILITY WAS REDUCED MORE BY THE NASA TECHNIQUE THAN BY SWAT



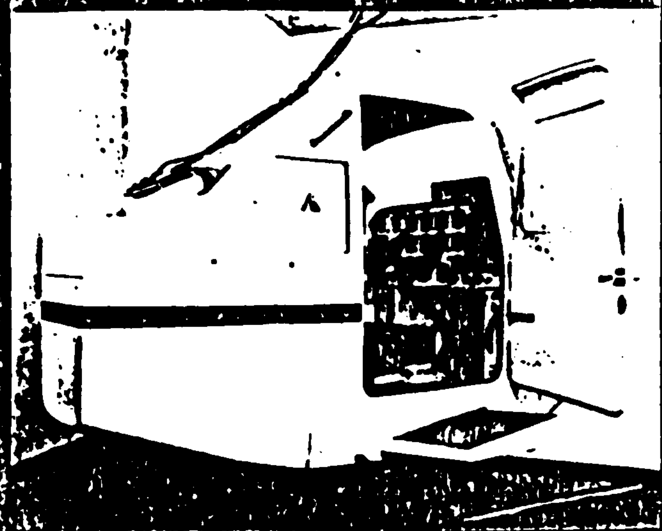


Vidulich & Tsang, 1985

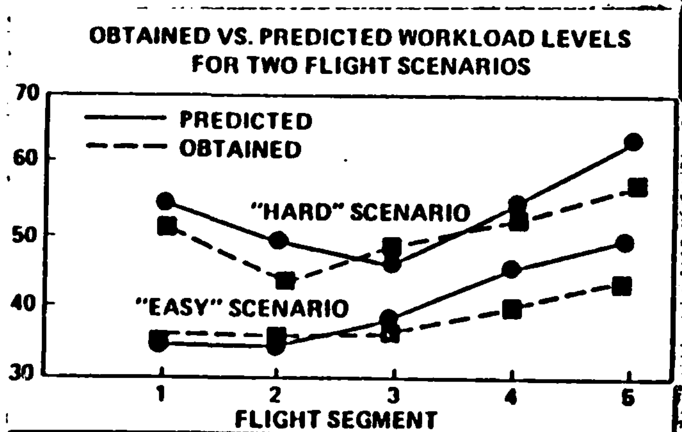
MEAN RATINGS AND STANDARD DEVIATIONS ACROSS
THE 35 TASK CONDITIONS FOR BOTH WORKLOAD GROUPS.

		NASA		SWAT	
SESSION	TASK	RATING	S.D.	RATING	S.D.
SESSION 4	LTR5	41	13.3	37	32.5
	RTR3	38	12.9	22	11.6
	RTR5	43	13.6	38	27.2
	RTR7	48	13.8	42	32.5
	RTRV	45	16.9	32	22.9
	VM	32	13.2	14	16.5
	VS	45	17.6	30	24.7
	AM	48	18.0	42	27.5
	AS	57	18.8	47	25.5
SESSION 7	LTR5	35	11.0	13	8.7
	RTR5	38	21.7	13	13.7
	RTRV	34	11.3	16	14.8
	LTR5-RTR5	46	14.7	49	23.3
	LTR5-RTRV	44	17.1	46	25.0
	VM-RTR5	44	15.2	48	23.2
	VS-RTR5	43	17.6	43	25.5
	AM-RTR5	50	15.9	55	25.0
	AS-RTR5	50	15.0	67	22.6
	VM-RTRV	40	18.0	43	26.1
	VS-RTRV	38	20.2	47	23.3
	AM-RTRV	52	16.2	55	24.0
	AS-RTRV	50	19.1	58	31.7
SESSION 10	LTR5	28	13.7	15	17.4
	RTR5	29	14.2	17	16.9
	RTRV	32	14.5	23	19.3
	LTR5-RTR5	43	16.5	67	29.4
	LTR5-RTRV	43	17.4	67	25.0
	VM-RTR5	46	15.9	57	28.9
	VS-RTR5	42	16.0	47	26.5
	AM-RTR5	50	18.4	57	31.1
	AS-RTR5	48	19.5	63	28.2
	VM-RTRV	46	15.8	49	24.3
	VS-RTRV	43	16.5	49	24.4
	AM-RTRV	51	18.6	64	22.0
	AS-RTRV	53	16.5	71	29.4

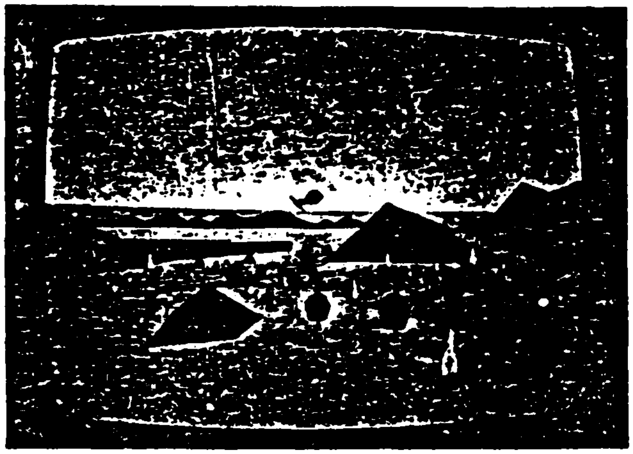
WORKLOAD PREDICTION



- OBJECTIVE
  - DEVELOP PREDICTIVE MODEL FOR CONSTRUCTING STANDARDIZED SIMULATION SCENARIOS
- APPROACH
  - CONDUCT PILOT OPINION SURVEYS
  - DEVELOP PREDICTIVE MODEL
  - TEST MODEL PREDICTIONS IN GAT SIMULATOR
- RESULTS:
  - OBJECTIVE AND SUBJECTIVE MEASURES OF PILOT WORKLOAD CLOSELY MATCH MODEL PREDICTIONS
- OPERATIONAL APPLICATIONS
  - HELICOPTER NOE FLIGHT
  - TRANSPORT OPERATIONS
  - SHUTTLE REMOTE MANIPULATOR ARM



HELICOPTER HUMAN FACTORS RESEARCH
SINGLE PILOT ADVANCED COCKPIT ENGINEERING SIMULATION



MISSION SCENARIO PROFILE

START AND END WAYPOINT
NAVIGATION
PREDETERMINED HOVER/FIRE POINT
HOVER AND BOB-UP MANEUVER
AIR TO AIR ENGAGEMENT
ELAPSED TIME/OR DISTANCE

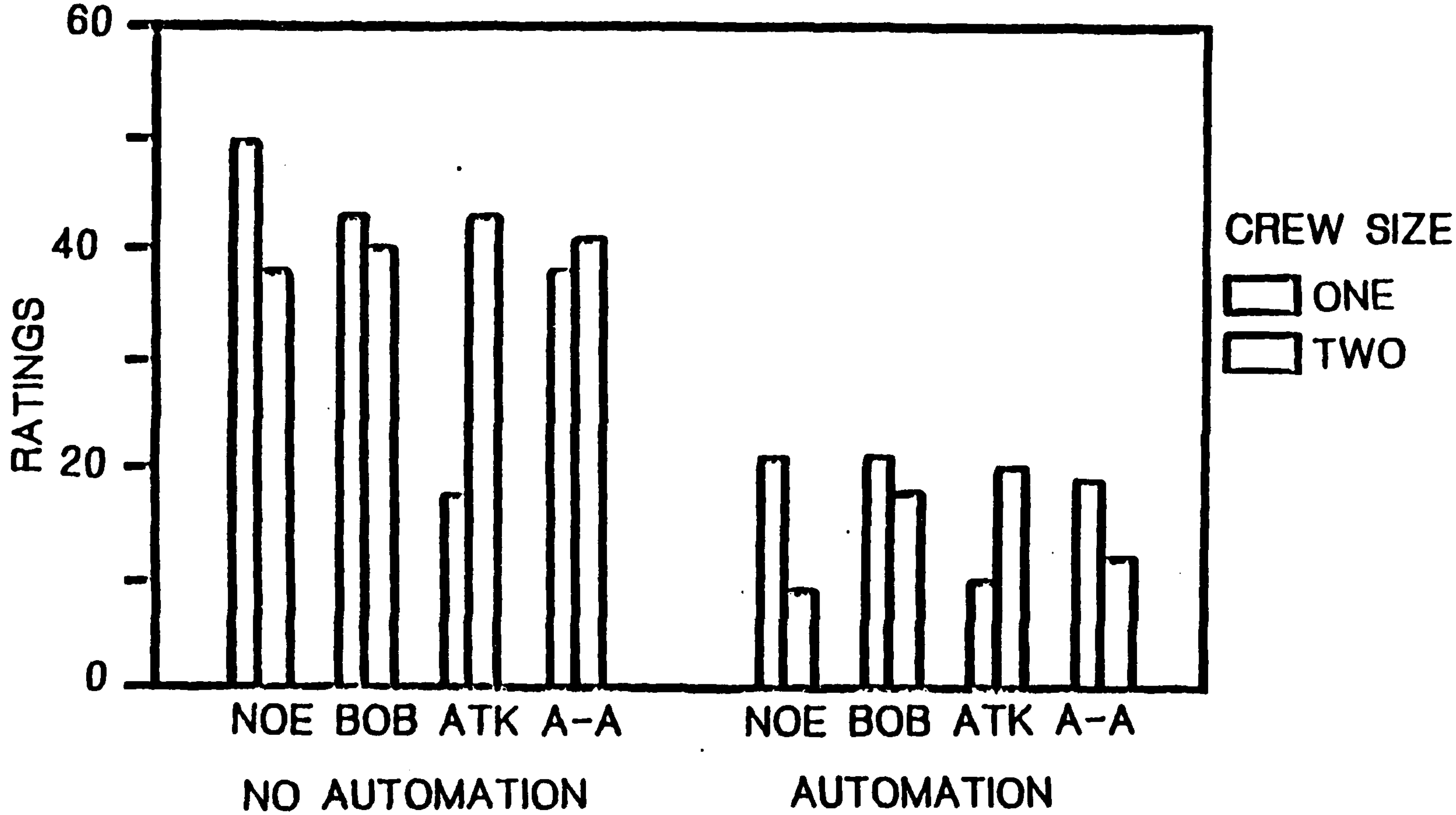
OBJECTIVE
- INVESTIGATE SINGLE-PILOT OPERATION IN THE NAP OF THE EARTH (NOE), FLIGHT COMBAT ENVIRONMENT

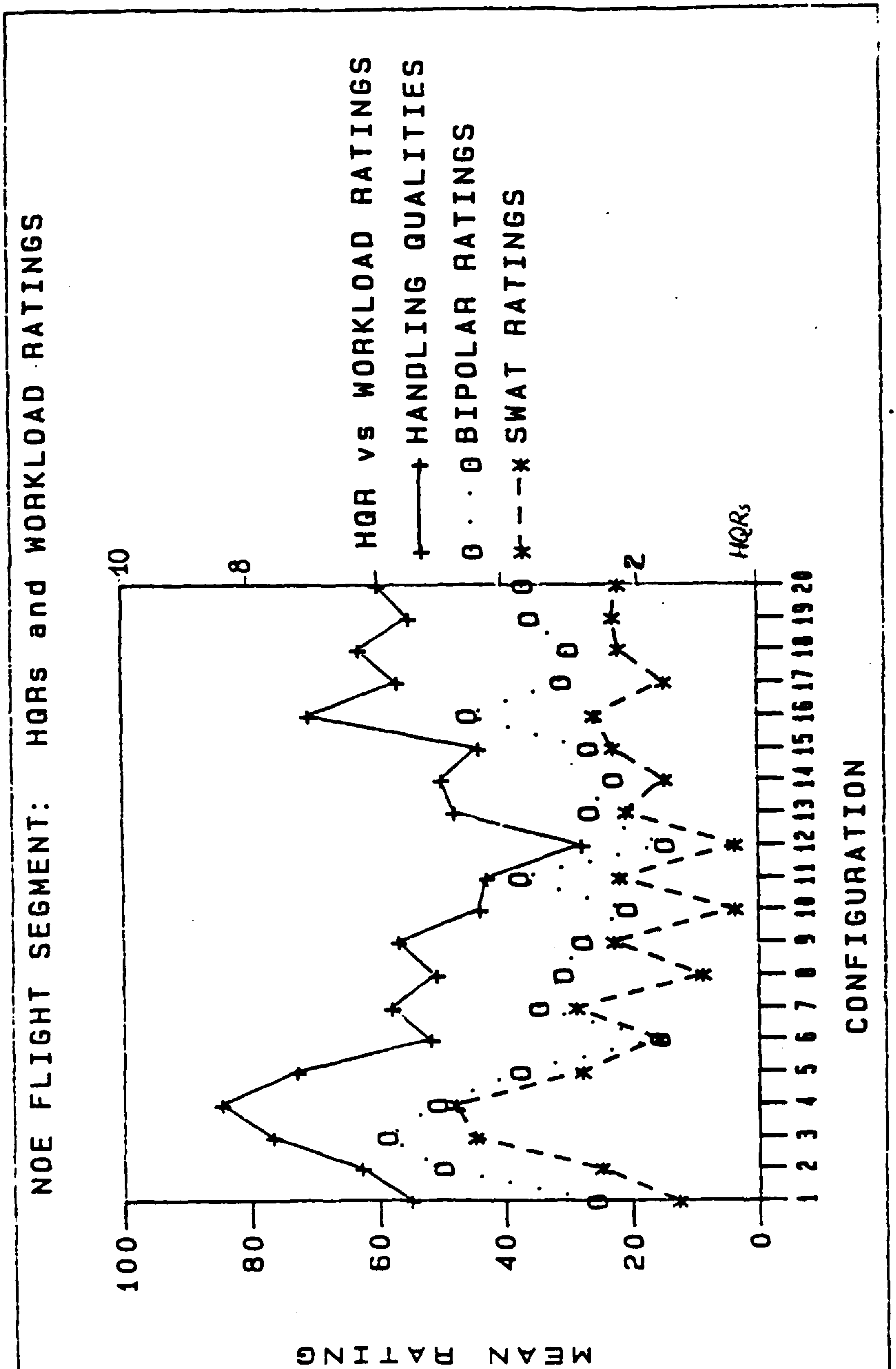
APPROACH
OBTAIN PILOT WORKLOAD, PERFORMANCE, AND HANDLING QUALITY RATINGS FOR DUAL AND SINGLE PILOT OPERATIONS

- SIMULATOR - AMES VERTICAL MOTION SIMULATOR
- AIRCRAFT MODEL - UH60 HELICOPTER
- COCKPIT CONTROLLERS - 2 + 1 + 1
- ADVANCED DIGITAL OPTICAL CONTROL SYSTEM (ADOCS)
  - ATTITUDE COMMAND/STABILIZATION OR HYBRID CONTROL SYSTEM
  - SELECTABLE HEADING, ALTITUDE, POSITION, AIRSPEED HOLD, TURN COORDINATION
- MISSION MANAGEMENT AND COMMUNICATIONS

RESULTS
- HIGH WORKLOAD FOR ONE PILOT
- ONLY ONE CONFIGURATION SATISFACTORY FOR NOE (ALTITUDE HOLD CRITICAL)
- HOVER HOLD ESSENTIAL FOR SINGLE PILOT OPERATIONS
- MISSION MANAGEMENT TASKS DEGRADED HANDLING QUALITY RATINGS

VMS CREW COMPLEMENT COMPARISON




NOE FLIGHT SEGMENT: HQRs and WORKLOAD RATINGS
HQR vs WORKLOAD RATINGS
HANDLING QUALITIES
BIPOLAR RATINGS
SWAT RATINGS
MEAN RATING
CONFIGURATION
HQRs
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
0 20 40 60 80 100
2 8 10

PHASE IV

- DEVELOP THE NASA TASK LOAD INDEX (NASA-TLX)
  - SELECT SIX SUBSCALES
  - TASK-RELATED WEIGHTING PROCEDURE
- EXPERIMENTALLY VALIDATE THE NASA TASK LOAD INDEX
  - 15 LABORATORY TASKS
  - SIMULATED FLIGHT
  - INFLIGHT
- APPLY VALIDATED TECHNIQUE
  - LABORATORY EXPERIMENTS
  - SUPERVISORY CONTROL SIMULATION
  - ARTI SIMULATION EVALUATIONS
  - SPACE STATION PROX-OPS DISPLAY EVALUATION

THE TYPES OF EXPERIMENTAL TASKS INCLUDED IN THE WORKLOAD RATING SCALE DEVELOPMENT EFFORT

- SIMPLE, COGNITIVELY-LOADING TASKS
  CHOICE REACTION TIME, MEMORY SEARCH, MENTAL ARITHMETIC, MENTAL ROTATION, PATTERN MATCH
- SIMPLE, MANUALLY-LOADING TASKS
  ONE AND TWO AXIS TRACKING
- CONCURRENT, INDEPENDENT DUAL-TASKS
  TRACKING + MEMORY SEARCH, MENTAL ROTATION
- SERIAL, INTEGRATED "FITTSBERG" TASKS
  TARGET ACQUISITION + MEMORY SEARCH, MENTAL ARITHMETIC, RHYMING, PATTERN MATCH, PREDICTION, TIME ESTIMATION
- COMPLEX SUPERVISORY CONTROL SIMULATIONS ("POPCORN")
- PART-TASK AND FULL-MISSION AIRCRAFT SIMULATIONS

SUBJECTIVE IMPORTANCE OF NINE FACTORS TO WORKLOAD

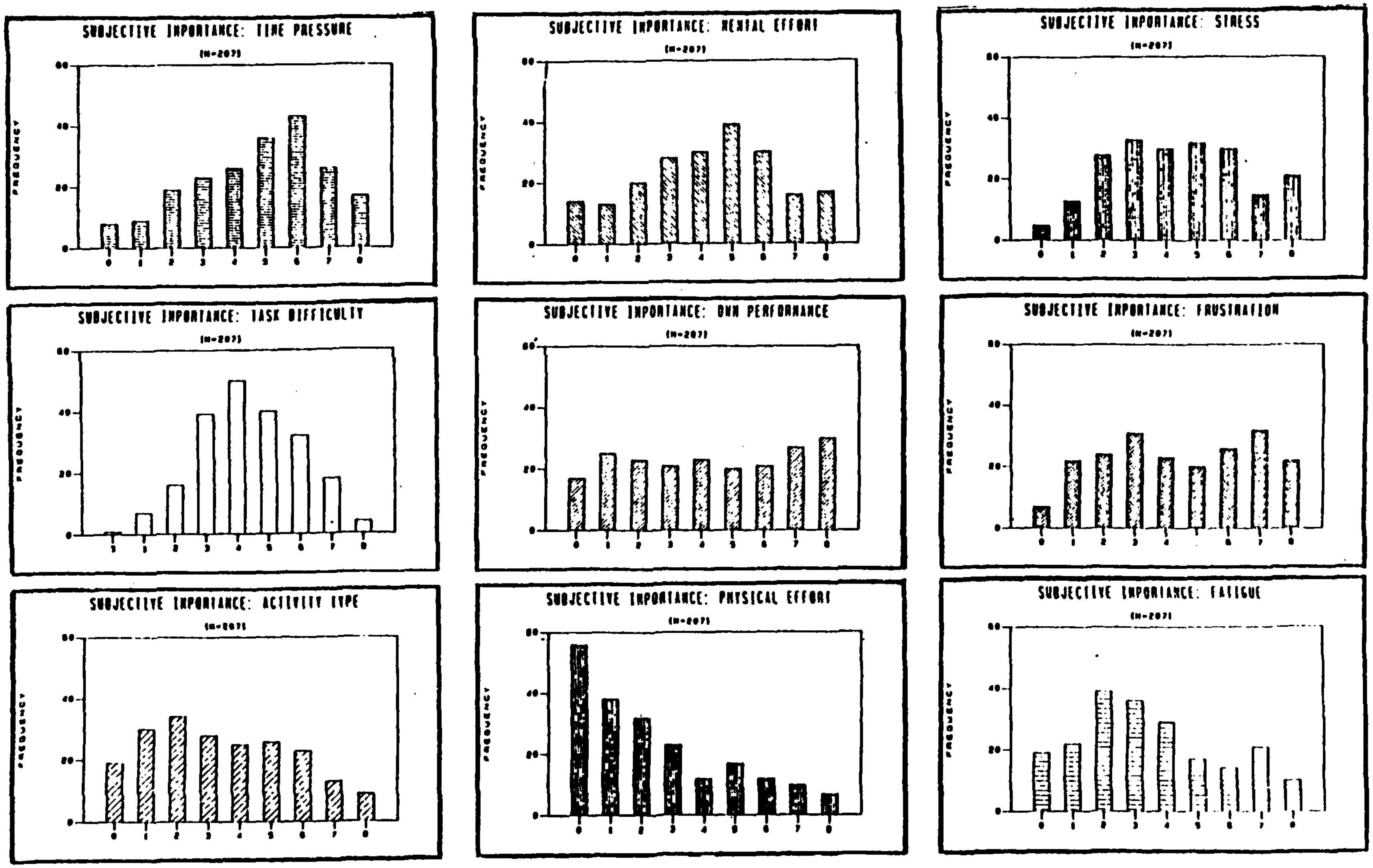


Table 3: POPULATION

Correlations among subjective importance values assigned to 9 workload-related factors

	TD	TP	OP	PE	ME	FR	ST	FA
TP	.05							
OP	-.08	-.24						
PE	-.12	-.31	-.07					
ME	.16	-.24	-.01	-.05				
FR	-.37	.05	-.21	-.26	-.30			
ST	-.21	.07	-.24	-.35	-.28	.32		
FA	-.21	-.03	-.46	.03	-.36	.10	.24	
AT	.08	-.17	.08	.17	.30	-.40	-.50	-.34

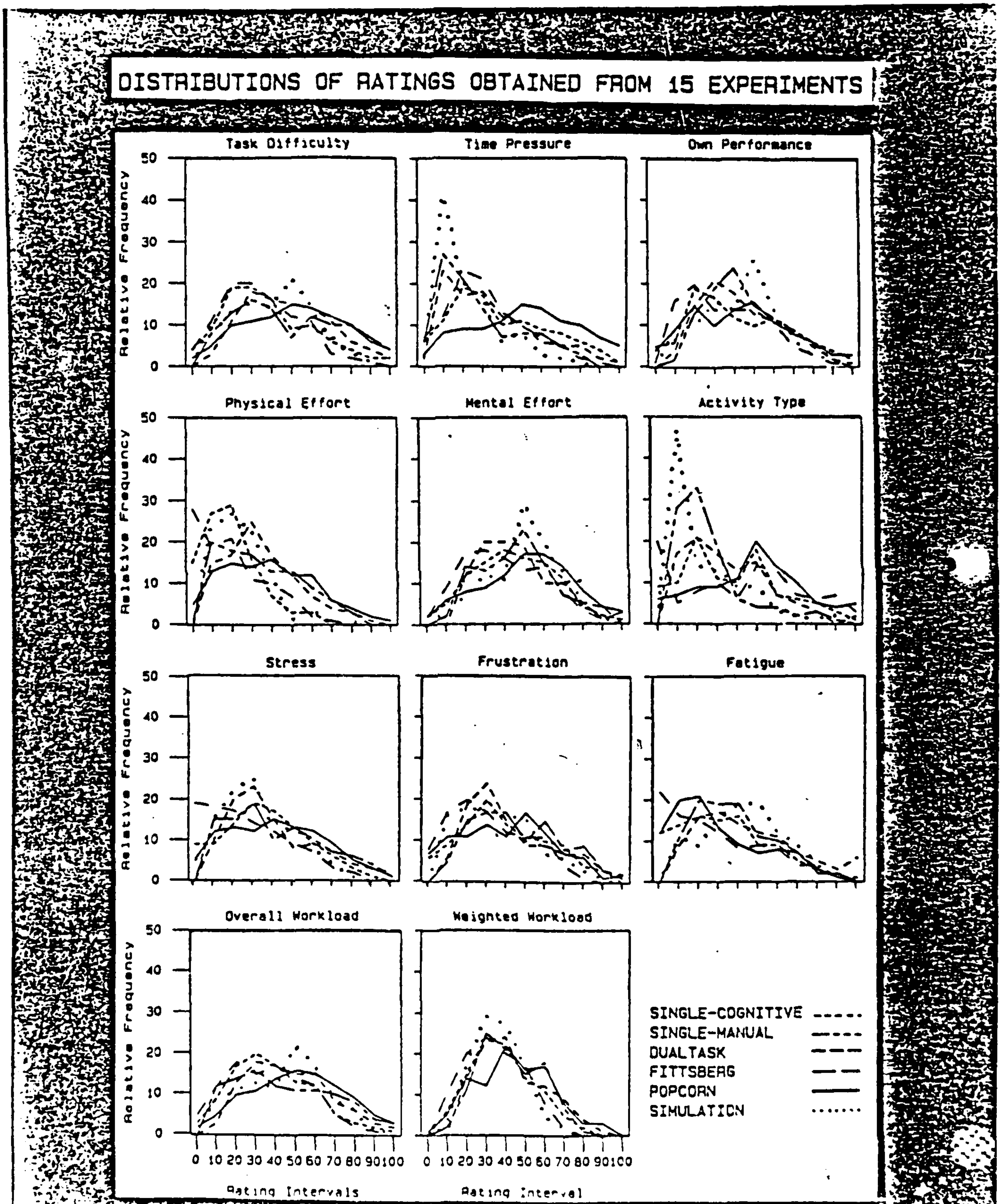

DISTRIBUTIONS OF RATINGS OBTAINED FROM 15 EXPERIMENTS
Task Difficulty
Time Pressure
Own Performance
Physical Effort
Mental Effort
Activity Type
Stress
Frustration
Fatigue
Overall Workload
Weighted Workload
Relative Frequency
0 10 20 30 40 50 60 70 80 90 100
Rating Intervals
Rating Interval
SINGLE-COGNITIVE
SINGLE-MANUAL
DUALTASK
FITTSBERG
POPCORN
SIMULATION

Figure 5.
SINGLE-COGNITIVE Category:
Summary of ratings
(Ns X Nc = 554).

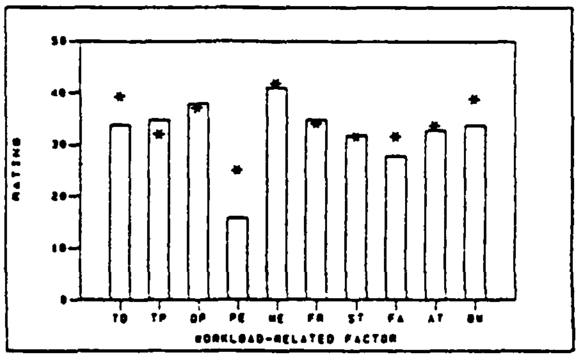


Figure 6.
SINGLE-MANUAL Category:
Summary of ratings
(Ns X Nc = 240).

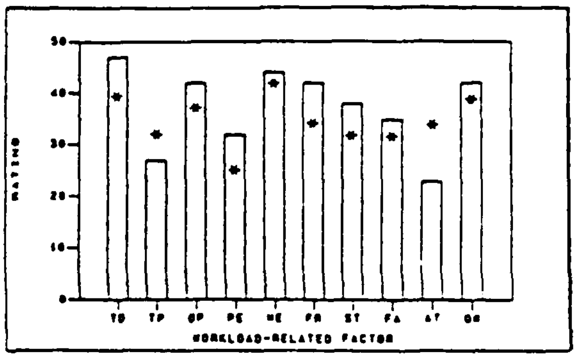


Figure 7.
DUAL-TASK Category:
Summary of ratings
(Ns X Nc = 732).

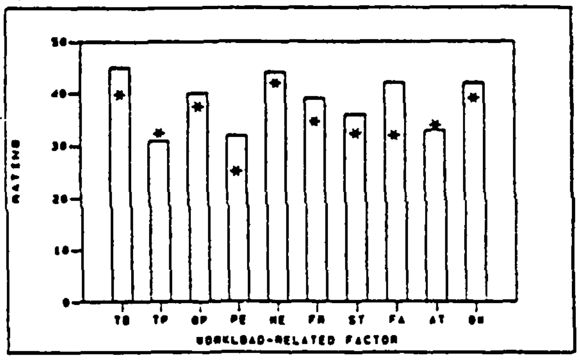


Figure 8.
FITTSBERG Category:
Summary of ratings
(Ns X Nc = 918).

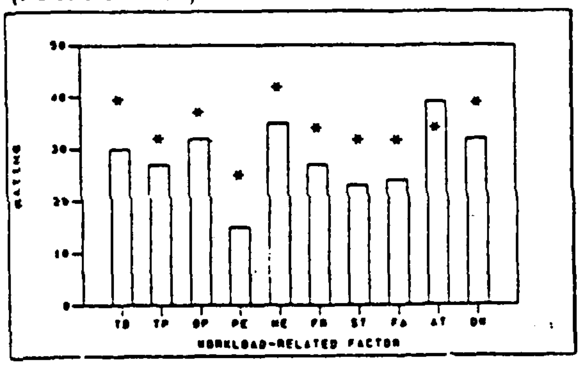


Figure 9.
POPCORN Category:
Summary of ratings
(Ns X Nc = 504).

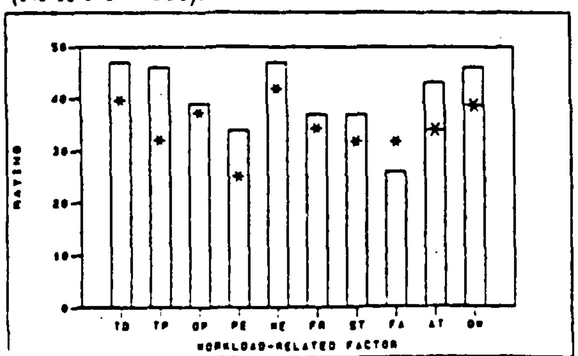
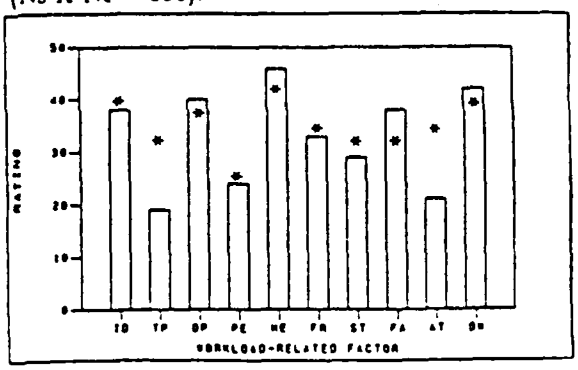


Figure 10.
SIMULATION Category:
Summary of ratings
(Ns X Nc = 396).

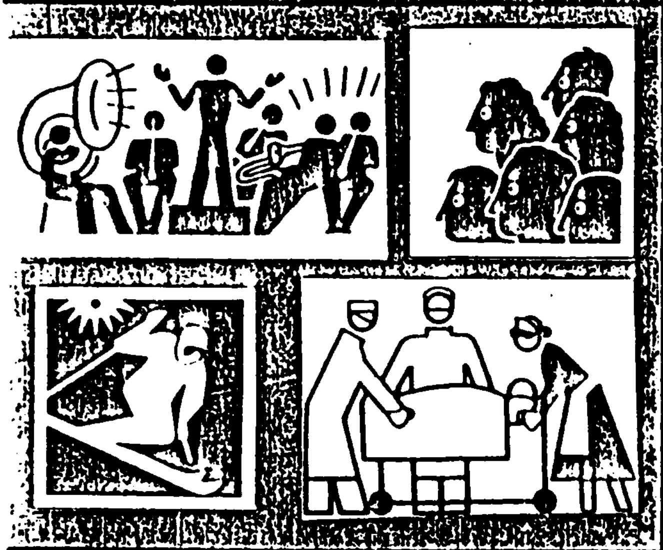


Correlations among raw bipolar ratings and WWL scores

	TD	TP	OP	PE	ME	FR	ST	FA	AT	OW
TP	.64									
OP	.58	.50								
PE	.53	.57	.38							
ME	.76	.58	.53	.47						
FR	.65	.60	.68	.45	.61					
ST	.63	.66	.48	.56	.60	.71				
FA	.38	.33	.40	.40	.37	.51	.52			
AT	.28	.29	.11	.20	.30	.21	.21	.11		
OW	.83	.60	.50	.52	.73	.63	.62	.40	.30	
WW	.83	.79	.71	.65	.79	.82	.80	.60	.39	.77

A priori rank-order of factors (weights) compared to empirical associations with OW ratings

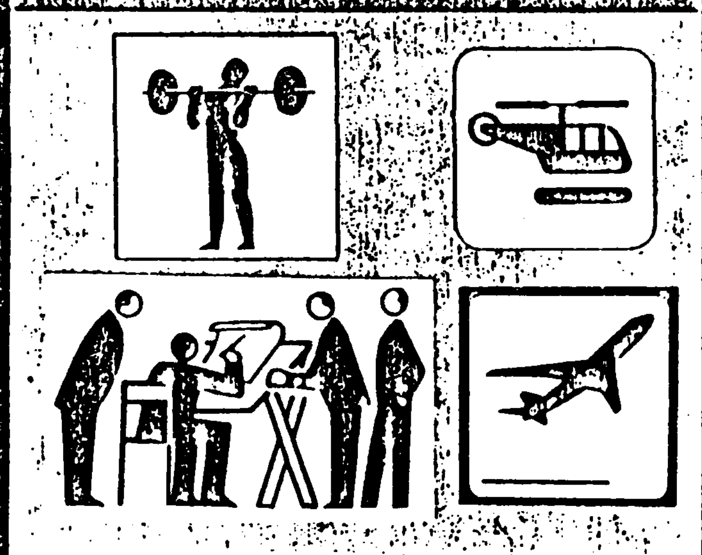
	Weight	Loading	Correlation with: OW
TP	4.75	.09	.60
TD	4.50	.55	.83
ME	4.36	.21	.73
OP	3.95	-.02	.50
ST	4.56	.10	.62
FR	4.51	.01	.63
FA	3.56	-.01	.40
AT	3.60	.01	.30
PE	2.21	.07	.52

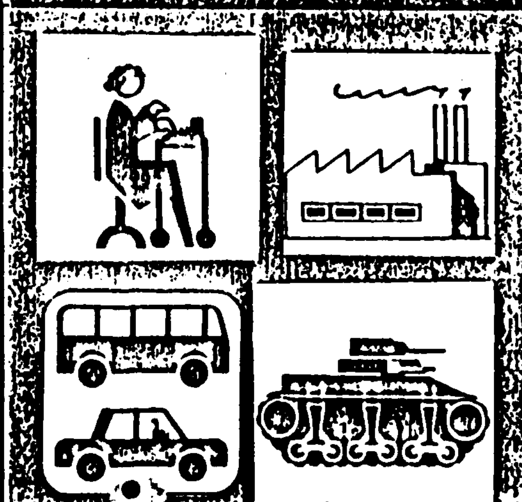


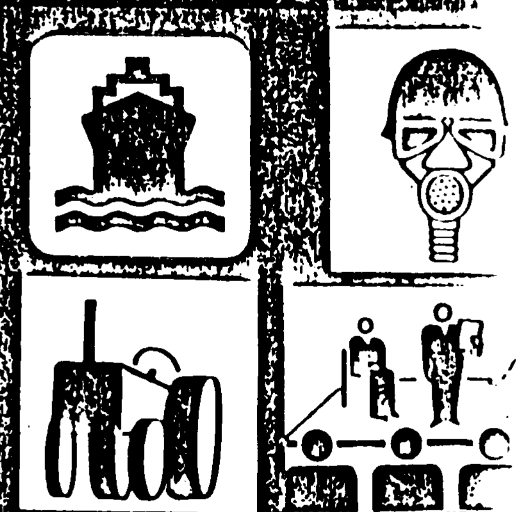
THE TERM "WORKLOAD" REPRESENTS A COLLECTION OF ATTRIBUTES THAT MAY OR MAY NOT BE RELEVANT IN A GIVEN TASK.

THE SUBEJECTIVE EXPERIENCE OF WORKLOAD EMERGES FROM THE INTERACTION BETWEEN OBJECTIVE TASK REQUIREMENTS AND AN INDIVIDUAL'S RESPONSE TO TI"

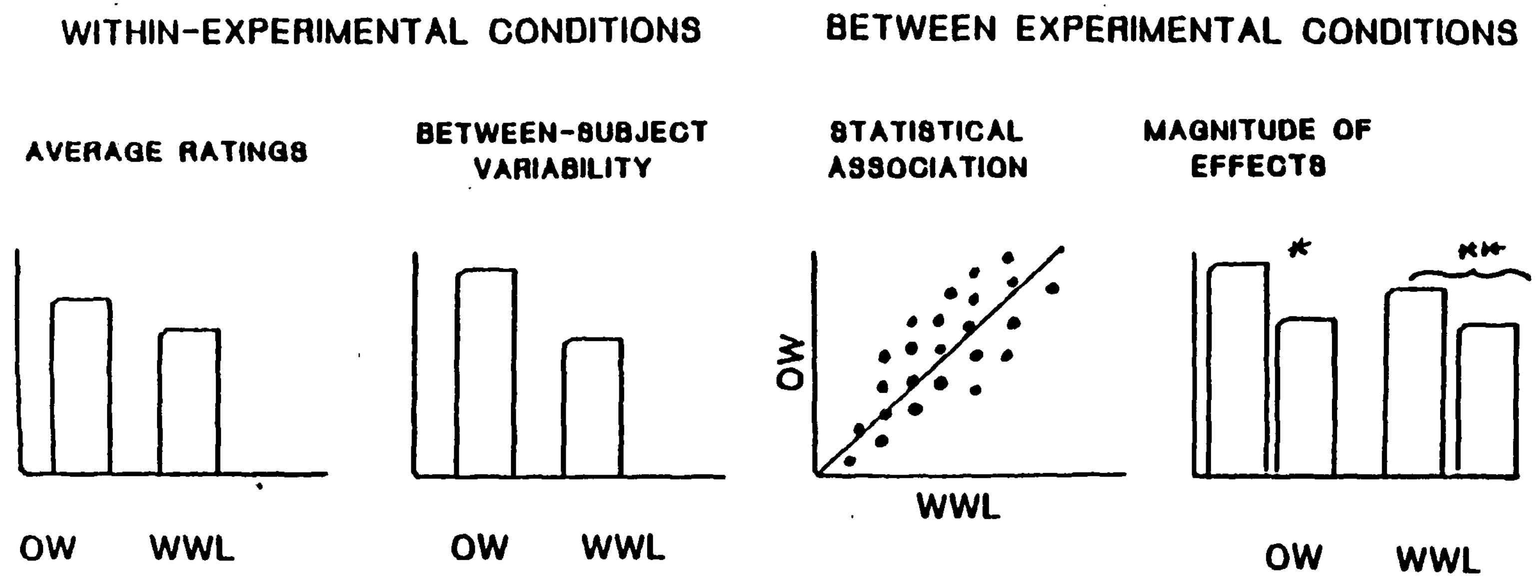
THUS, WORKLOAD IS NOT AN OBJECTIVE ENTITY AND ITS SOURCES AND CONSEQUENCES VARY ACROSS TASKS.







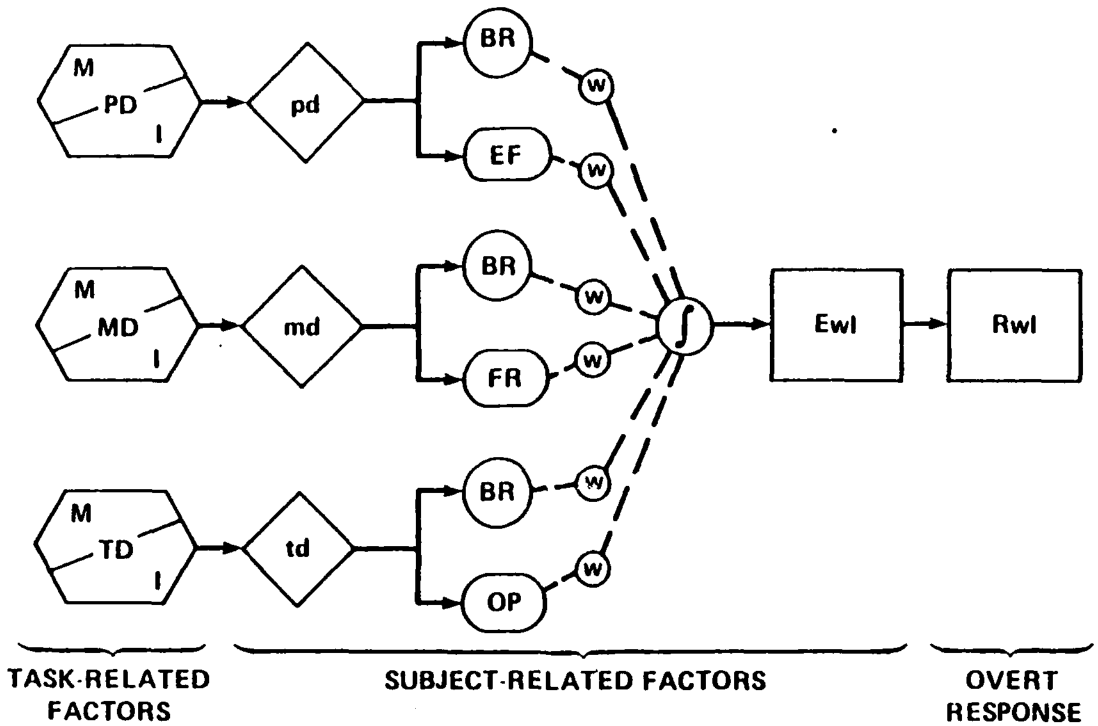
COMPARISON BETWEEN OVERALL WORKLOAD RATINGS AND WEIGHTED WORKLOAD SCORE



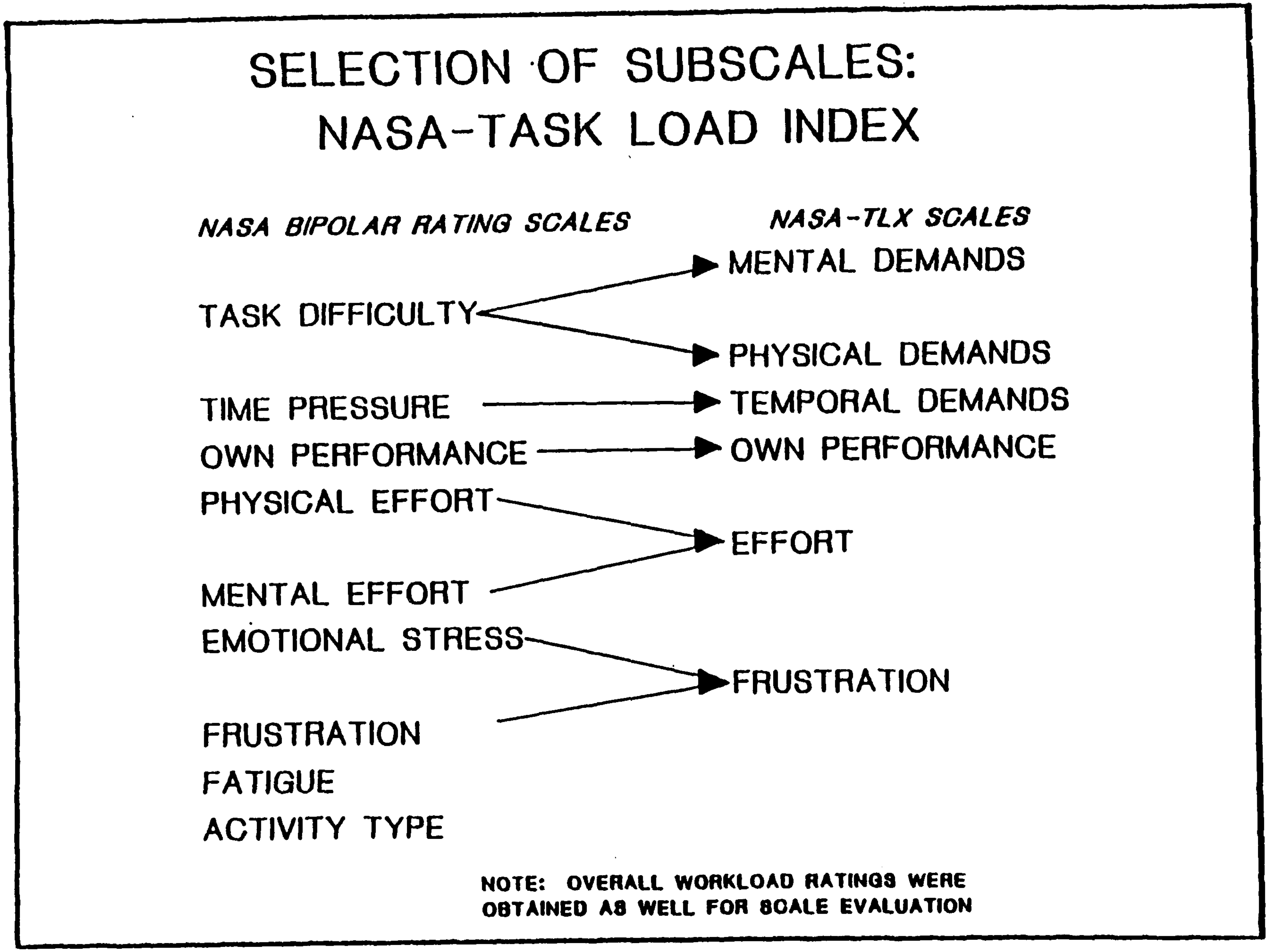
SUMMARY OF NASA BIPOLAR RATING SCALE EVALUATION

- SOURCES OF WORKLOAD VARY BETWEEN DIFFERENT TYPES OF TASKS
- RATINGS OF COMPONENT FACTORS ARE MORE DIAGNOSTIC THAN GLOBAL WORKLOAD RATINGS
- SUBJECTIVE WORKLOAD DEFINITIONS DO VARY (THEREBY CONTRIBUTING TO BETWEEN-SUBJECT VARIABILITY)
- HOWEVER, THEIR A PRIORI BIASES ABOUT WORKLOAD ARE UNRELATED TO THEIR RATINGS OF WORKLOAD AND WORKLOAD COMPONENTS
- SUBSCALES:
  - o SOME WERE HIGHLY CORRELATED (E.G., STRESS, FRUSTRATION)
    - THEY WERE COMBINED
  - o OTHERS WERE UNRELATED TO WORKLOAD (E.G., FATIGUE)
    - THEY WERE DELETED
  - o OTHERS WERE TOO BROAD (E.G., TASK DIFFICULTY)
    - THEY WERE SUBDIVIDED

MODEL OF SUBJECTIVE WORKLOAD ESTIMATION PROCESS



PD, MD, TD	*Objective physical, mental and temporal task demands*
M, I	Objective magnitudes and importance of sources of demands
pd, md, td	Psychological representations of task demands
BR	Behavioral responses to task demands
OP, EF, FR	Subjective responses/evaluations of behavioral responses
w	*Subjective weighting of factors*
Ewl	Integrated subjective experience of workload
Rwl	Formal numeric or verbal evaluation of workload


SELECTION OF SUBSCALES:
NASA-TASK LOAD INDEX
NASA BIPOLAR RATING SCALES
NASA-TLX SCALES
TASK DIFFICULTY
MENTAL DEMANDS
PHYSICAL DEMANDS
TIME PRESSURE
TEMPORAL DEMANDS
OWN PERFORMANCE
OWN PERFORMANCE
PHYSICAL EFFORT
MENTAL EFFORT
EFFORT
EMOTIONAL STRESS
FRUSTRATION
FRUSTRATION
FATIGUE
ACTIVITY TYPE
NOTE: OVERALL WORKLOAD RATINGS WERE
OBTAINED AS WELL FOR SCALE EVALUATION

RATING SCALE DEFINITIONS

Title	Endpoints	Descriptions
MENTAL DEMAND	*Low/High*	How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?
PHYSICAL DEMAND	*Low/High*	How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
TEMPORAL DEMAND	*Low/High*	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
PERFORMANCE	*Perfect/Failure*	How successful do you think you were in accomplishing the goals of the task set by the *experimenter* (or yourself)? How satisfied were you with your performance in accomplishing these goals?
EFFORT	*Low/High*	How hard did you have to work (mentally and physically) to accomplish your level of performance?
FRUSTRATION LEVEL	*Low/High*	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

Effort or Performance	Temporal Demand or Frustration
Temporal Demand or Effort	Ph[illegible]emand or Frustration
Performance or Frustration	Physical Demand or Temporal Demand
Physical Demand or Performance	Temporal Demand or Mental Demand

RATING SHEET

MENTAL DEMAND

Low High

PHYSICAL DEMAND

Low High

TEMPORAL DEMAND

Low High

PERFORMANCE

Good Poor

EFFORT

Low High

FRUSTRATION

Low High

OVERALL WORKLOAD

Low High

SAMPLE APPLICATION OF NASA WORKLOAD SCALE

EXAMPLE:

COMPARE WORKLOAD OF TWO TASKS THAT REQUIRE A SERIES OF DISCRETE RESPONSES. THE PRIMARY DIFFICULTY MANIPULATION IS THE INTER-STIMULUS INTERVAL (ISI) – (TASK 1 = 500 msec. TASK 2 = 300 msec)

PAIR-WISE COMPARISONS OF FACTORS:

INSTRUCTIONS: SELECT THE MEMBER OF EACH PAIR THAT PROVIDED THE MOST SIGNIFICANT SOURCE OF WORKLOAD VARIATION IN THESE TASKS

PD / MD	TD / PD	TD / FR
TD / MD	OP / PD	TD / EF
OP / MD	FR / PD	OP / FR
FR / MD	EF / PD	OP / EF
EF / MD	TD / OP	EF / FR

TALLY OF IMPORTANCE SELECTIONS

MD III	= 3
PD	= 0
TD IIIII	= 5
OP I	= 1
FR III	= 3
EF III	= 3
SUM	= 15

RATING SCALES:

INSTRUCTIONS: PLACE A MARK ON EACH SCALE THAT REPRESENTS THE MAGNITUDE OF EACH FACTOR IN THE TASK YOU JUST PERFORMED

DEMANDS	RATINGS FOR TASK 1:			RATING		WEIGHT		PRODUCT
MD	LOW	I x I	HIGH	30	X	3	=	90
PD	LOW	I x I	HIGH	15	X	0	=	0
TD	LOW	I x I	HIGH	60	X	5	=	150
OP	EXCL	I x I	POOR	40	X	1	=	40
FR	LOW	I x I	HIGH	30	X	3	=	90
EF	LOW	I x I	HIGH	40	X	3	=	120
				SUM			=	490
				WEIGHTS (TOTAL)			=	15
				MEAN WWL SCORE			=	32

DEMANDS	RATINGS FOR TASK 2:			RATING		WEIGHT		PRODUCT
MD	LOW	I x I	HIGH	30	X	3	=	90
PD	LOW	I x I	HIGH	25	X	0	=	0
TD	LOW	I x I	HIGH	70	X	5	=	350
OP	EXCL	I x I	POOR	50	X	1	=	50
FR	LOW	I x I	HIGH	50	X	3	=	150
EF	LOW	I x I	HIGH	30	X	3	=	90
				SUM			=	730
				WEIGHTS (TOTAL)			=	15
				MEAN WWL SCORE			=	49

RESULTS:

SUBSCALES PINPOINT SPECIFIC SOURCE OF WORKLOAD VARIATION BETWEEN TASKS (TD) THE WWL SCORE REFLECTS THE IMPORTANCE OF THIS AND OTHER FACTORS AS WORKLOAD-DRIVERS AND THEIR SUBJECTIVE MAGNITUDE IN EACH TASK

CIRCADIAN CHANGES IN WORKLOAD, PERFORMANCE, AND PHYSIOLOGICAL RESPONSES

OBJECTIVE: TO DETERMINE WHETHER CHANGES IN SUBJECTIVE, BEHAVIORAL, AND PHYSIOLOGICAL MEASURES OF WORKLOAD ACCOMPANY THE WELL-KNOWN VARIATIONS IN BODY TEMPERATURE AND TASK PERFORMANCE THROUGHOUT THE DAY.

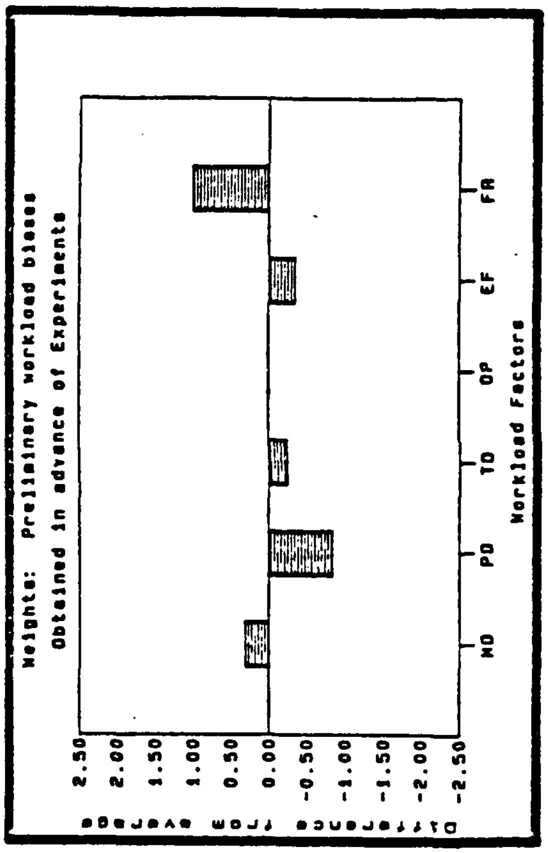
APPROACH: OBTAIN A VARIETY OF MEASURES FROM SIX MALE SUBJECTS DURING THEIR PERFORMANCE OF DIFFERENT TASKS WHILE THEY ARE CONFINED IN THE NASA HUMAN RESEARCH FACILITY FOR SEVEN DAYS.

EXPERIMENTAL TASKS:

- TARGET ACQUISITION
- CONTINUOUS MANUAL CONTROL
- STEP TRACKING
- ICONIC MEMORY
- SHORT-TERM MEMORY
- GRAMMATICAL REASONING
- TIME ESTIMATION
- VISUAL TARGET RECOGNITION
- MENTAL ROTATION
- MENTAL ARITHMETIC

EXPERIMENTAL MEASURES:

- BODY TEMPERATURE
- SUBJECTIVE RATINGS
- HEART RATE
- RESPONSE ACCURACY, SPEED


Weights: Preliminary workload biases
Obtained in advance of Experiments
Difference from average
2.50
2.00
1.50
1.00
0.50
0.00
-0.50
-1.00
-1.50
-2.00
-2.50
WO
PD
TD
OP
EF
FR
Workload Factors

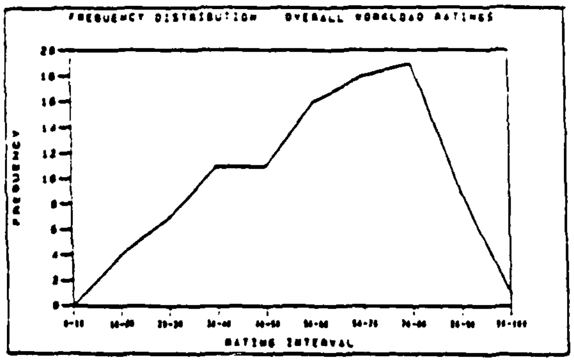
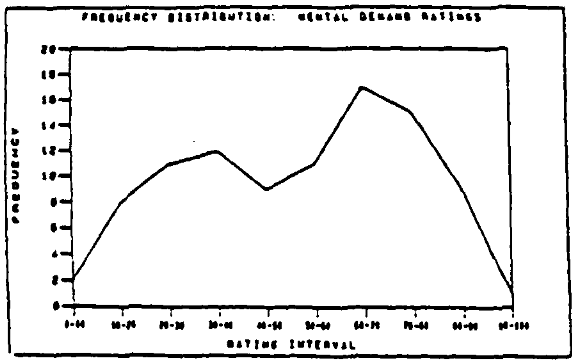
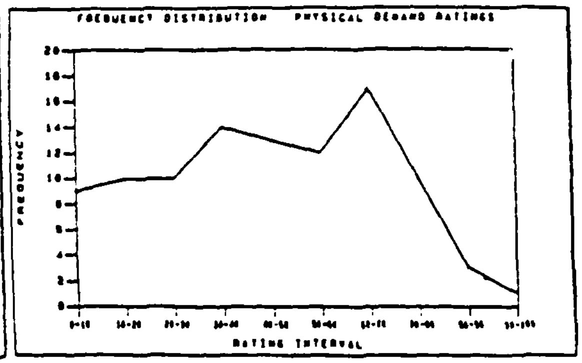
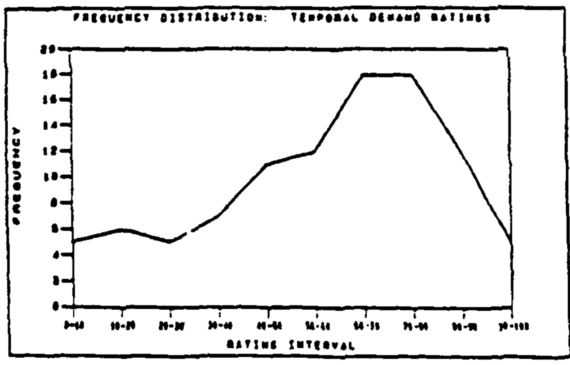
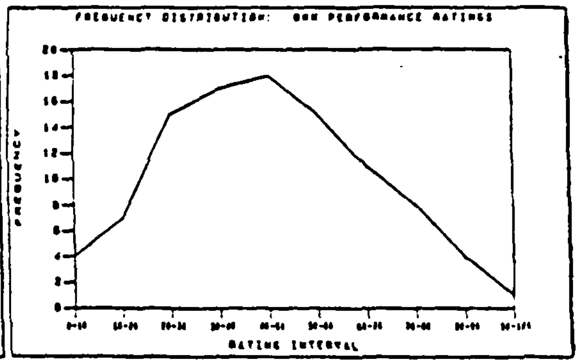


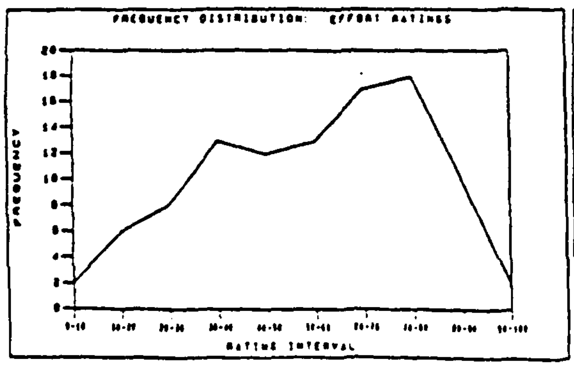
Figure 14. Distribution of ratings on six subscales obtained in the validation study. (Ns X Nc = 984).

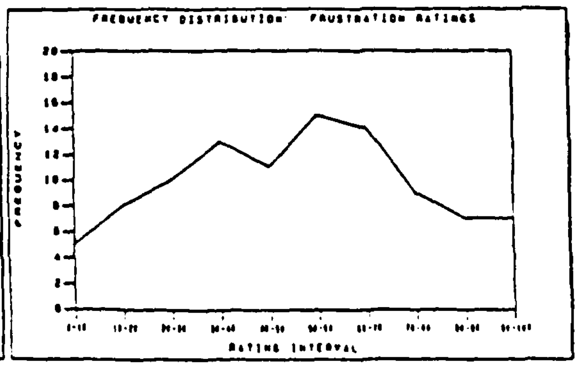




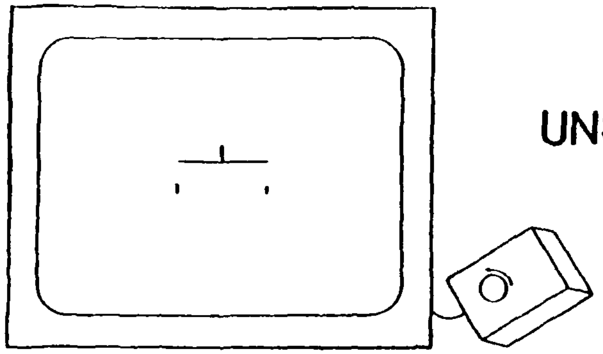




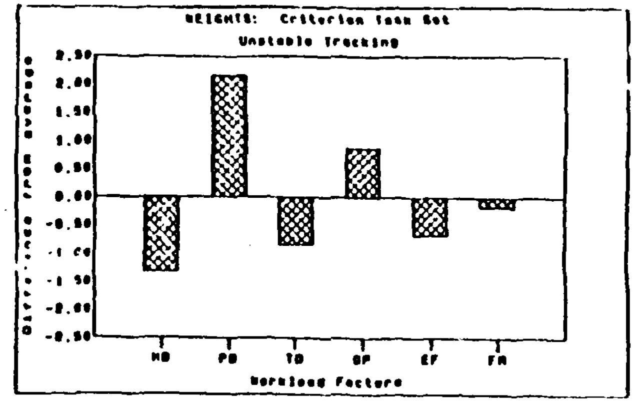


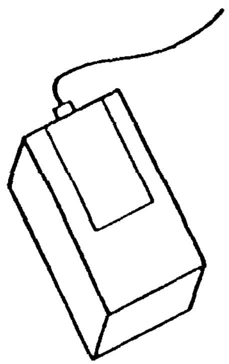


VALIDATION STUDY: CRITERION TASK SET

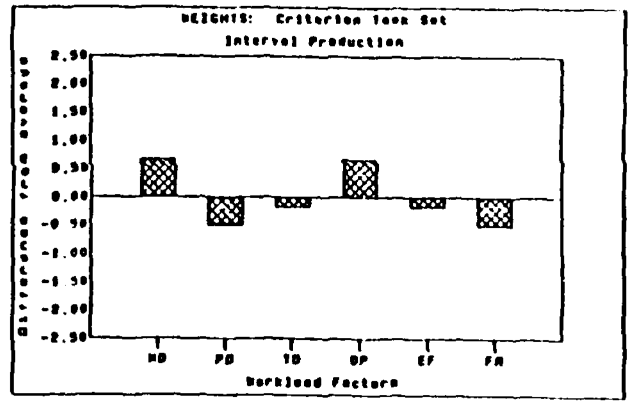


UNSTABLE TRACKING

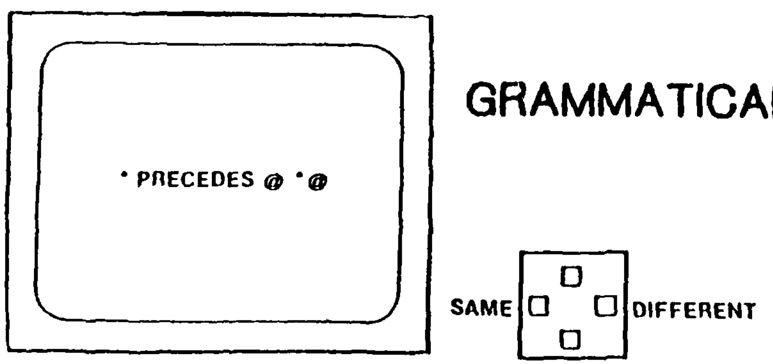


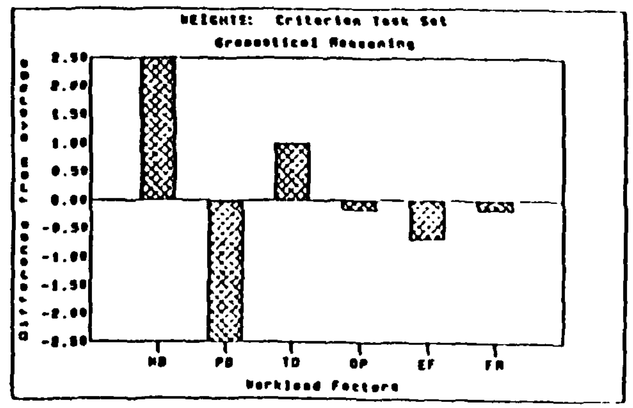


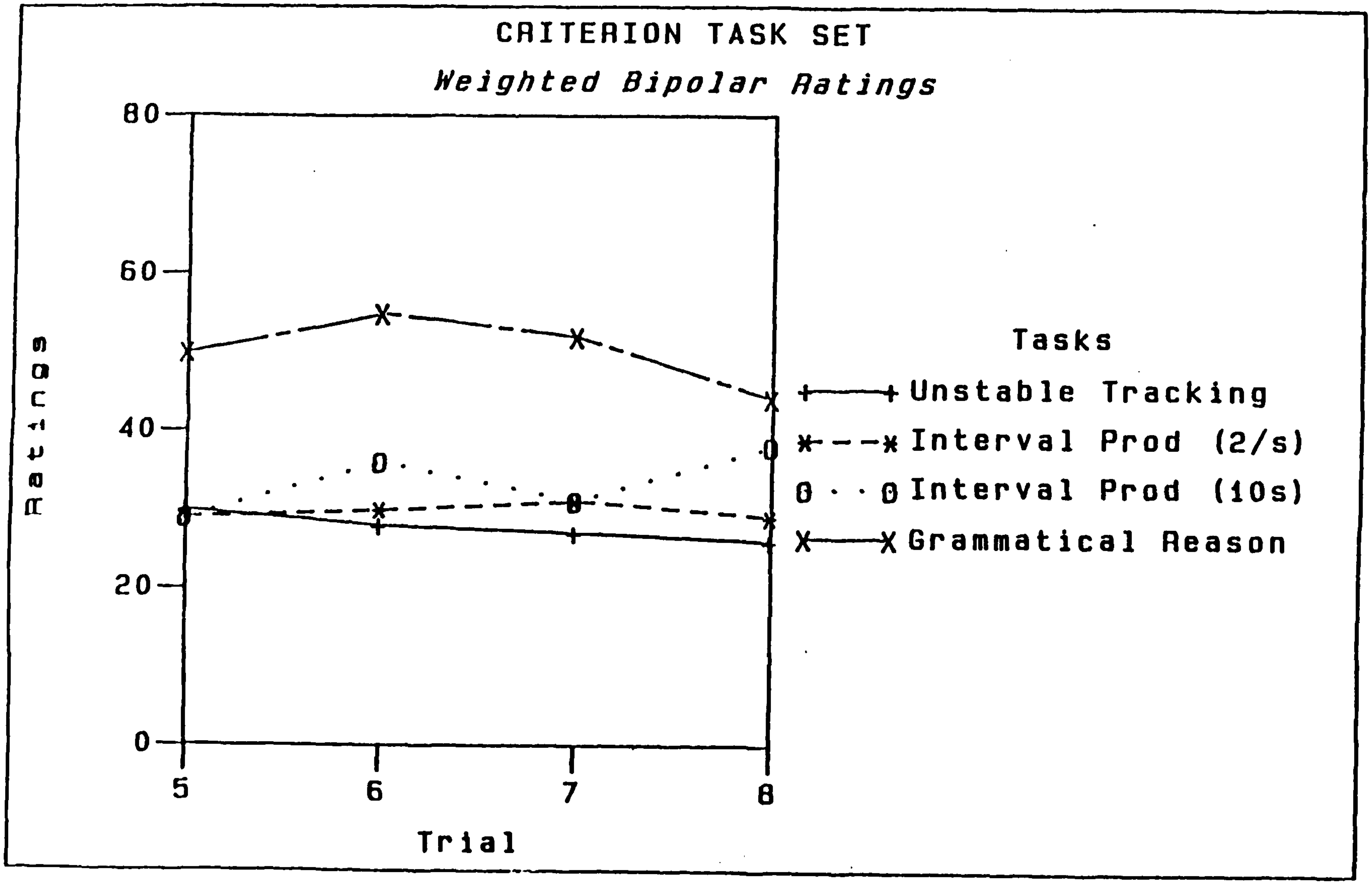
TIME ESTIMATION



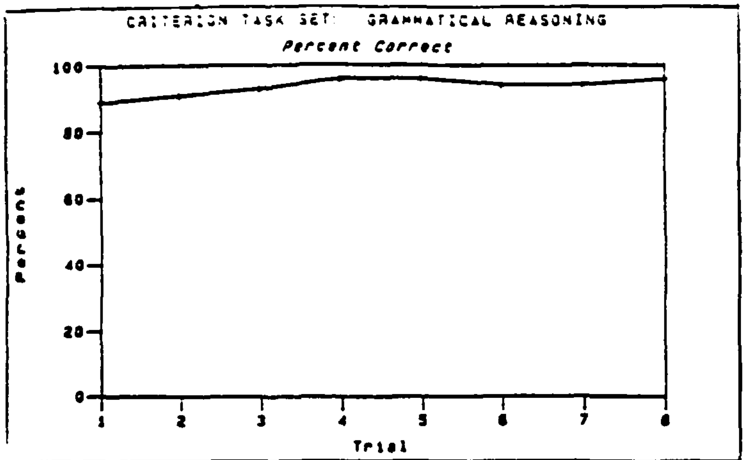
GRAMMATICAL REASONING

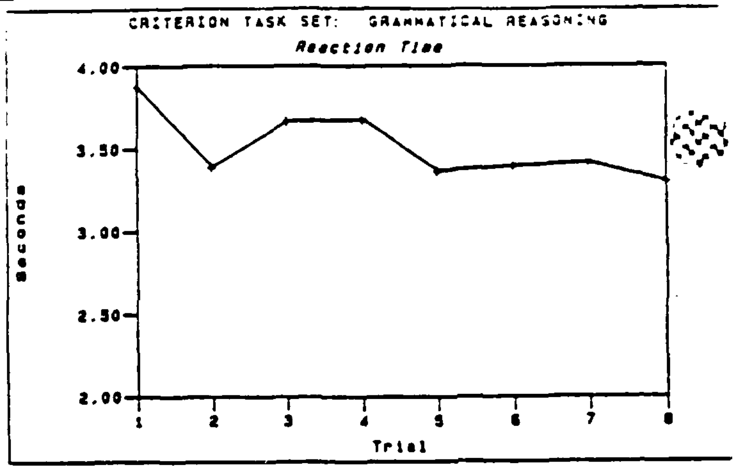


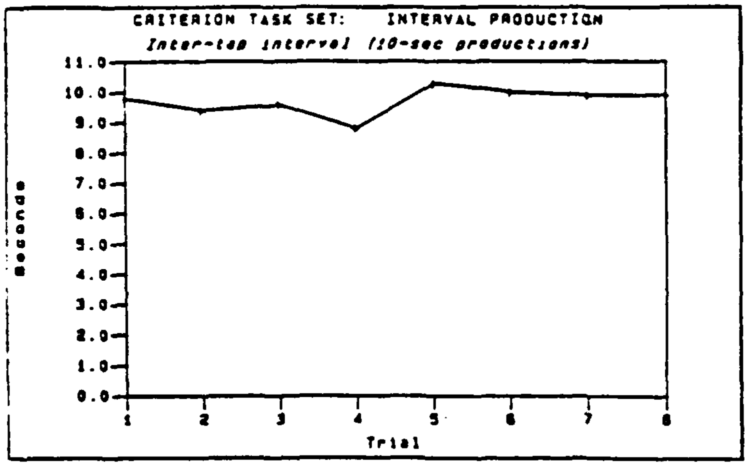


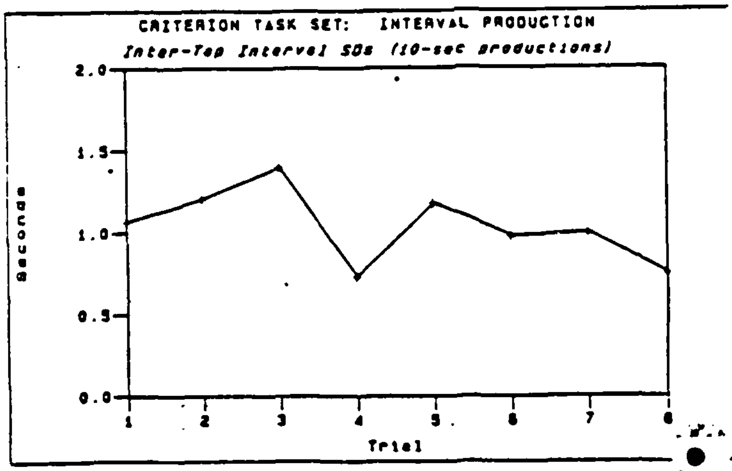


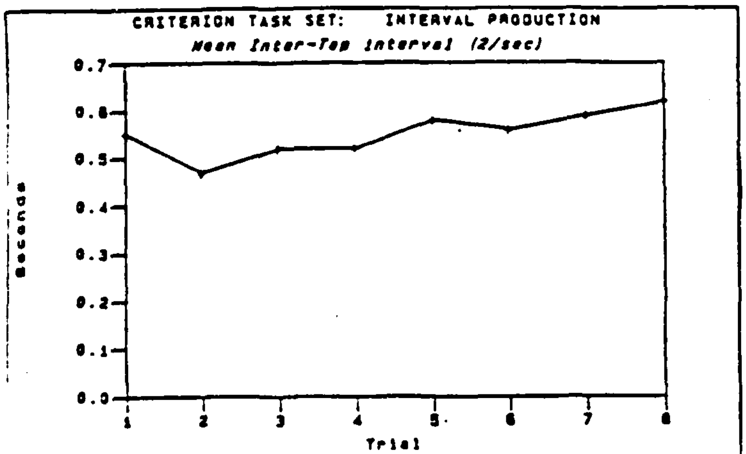
CRITERION TASK SET
Weighted Bipolar Ratings
Ratings
80
60
40
20
0
5
6
7
8
Trial
Tasks
Unstable Tracking
Interval Prod (2/s)
Interval Prod (10s)
Grammatical Reason

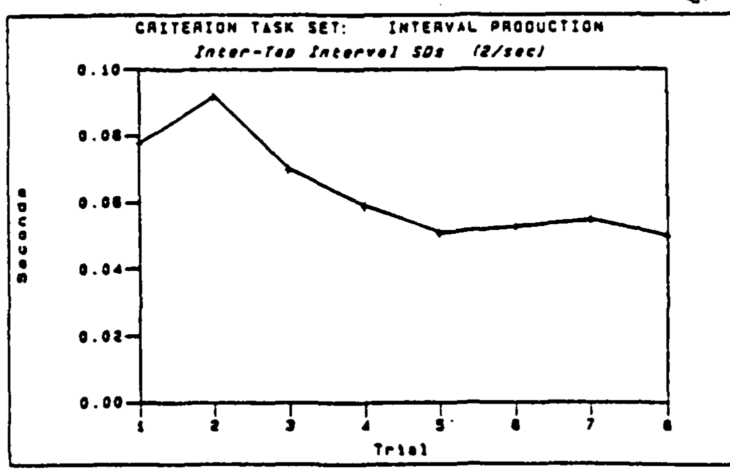

CRITERION TASK SET: GRAMMATICAL REASONING
Percent Correct
Percent
Trial
100
80
60
40
20
0
1 2 3 4 5 6 7 8

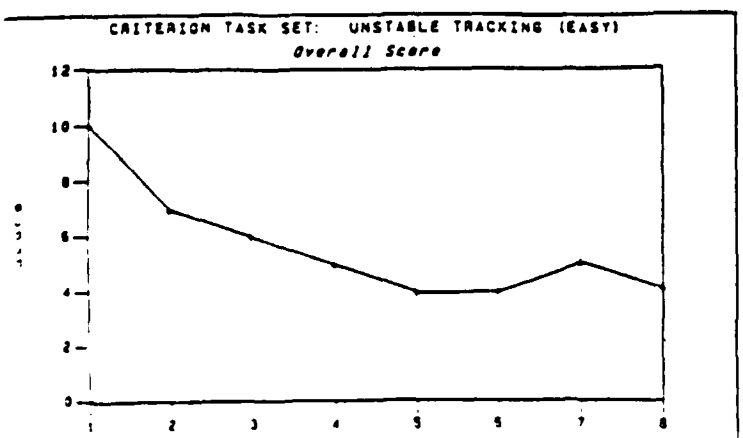

CRITERION TASK SET: GRAMMATICAL REASONING
Reaction Time
Seconds
Trial
4.00
3.50
3.00
2.50
2.00
1 2 3 4 5 6 7 8

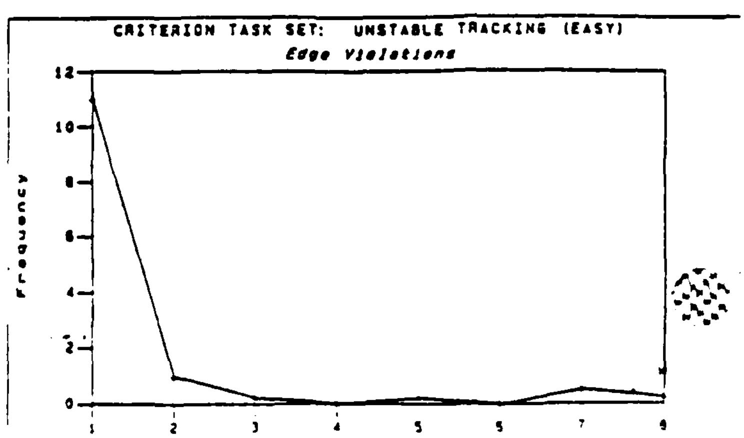

CRITERION TASK SET: INTERVAL PRODUCTION
Inter-tap interval (10-sec productions)
Seconds
Trial
11.0
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
1 2 3 4 5 6 7 8


CRITERION TASK SET: INTERVAL PRODUCTION
Inter-Tap Interval SDs (10-sec productions)
Seconds
Trial
2.0
1.5
1.0
0.5
0.0
1 2 3 4 5 6 7 8

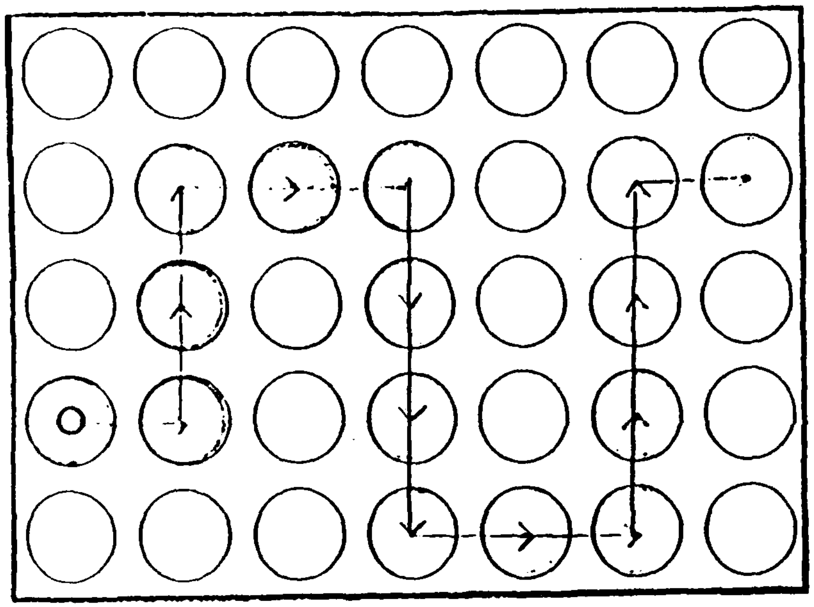

CRITERION TASK SET: INTERVAL PRODUCTION
Mean Inter-Tap Interval (2/sec)
Seconds
Trial
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
1 2 3 4 5 6 7 8

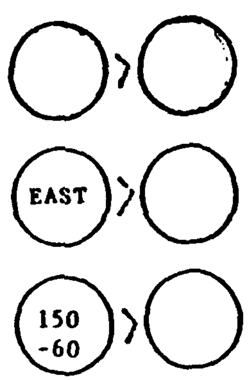

CRITERION TASK SET: INTERVAL PRODUCTION
Inter-Tap Interval SDs (2/sec)
Seconds
Trial
0.10
0.08
0.06
0.04
0.02
0.00
1 2 3 4 5 6 7 8

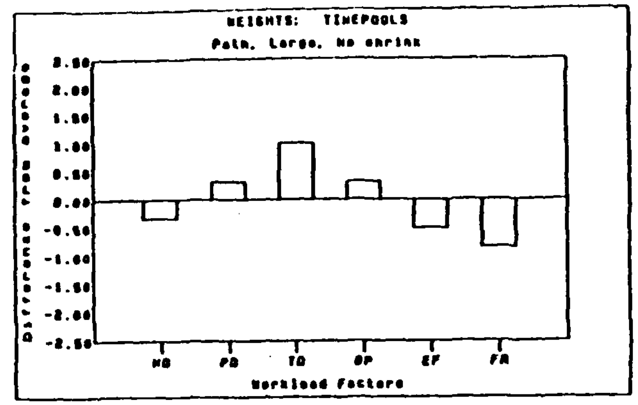

CRITERION TASK SET: UNSTABLE TRACKING (EASY)
Overall Score
12
10
8
6
4
2
0
1 2 3 4 5 6 7 8

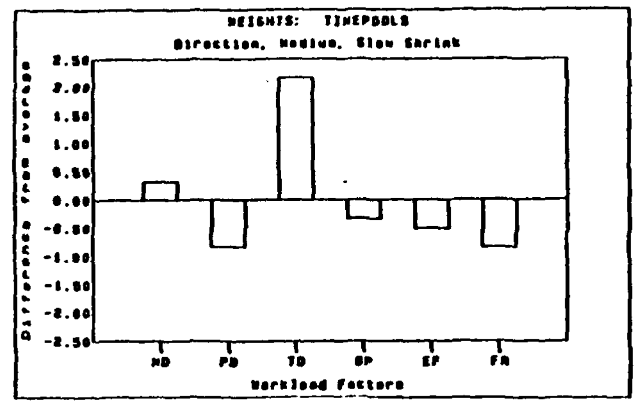

CRITERION TASK SET: UNSTABLE TRACKING (EASY)
Edge Violations
Frequency
12
10
8
6
4
2
0
1 2 3 4 5 6 7 8

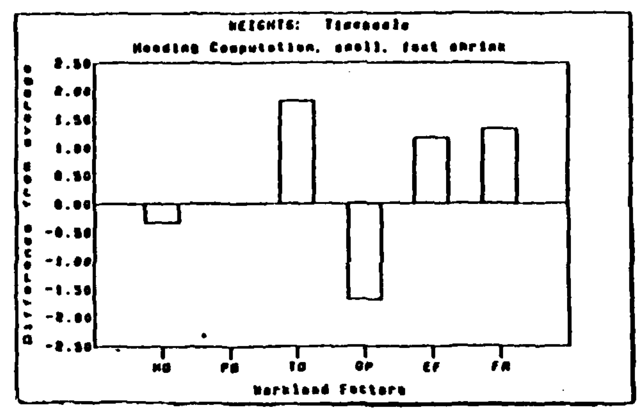
VALIDATION STUDY: TIMEPOOLS EXPERIMENT



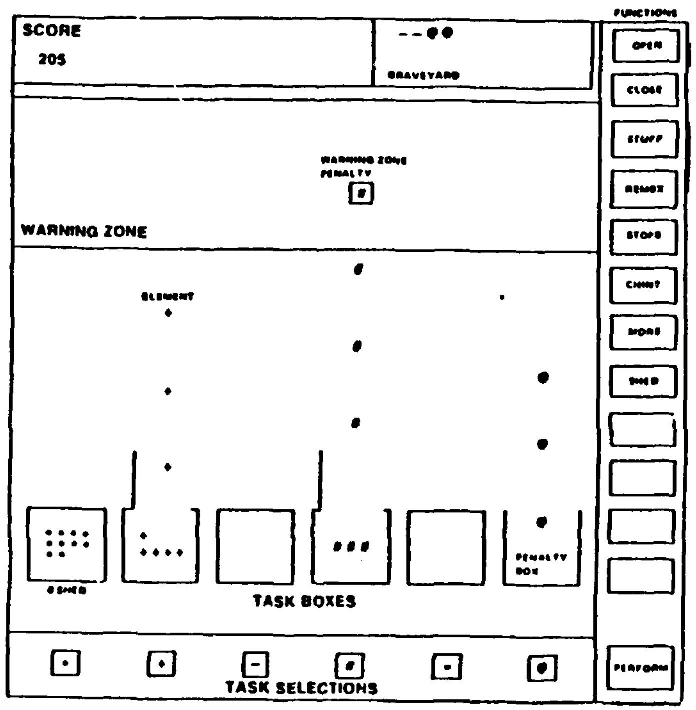


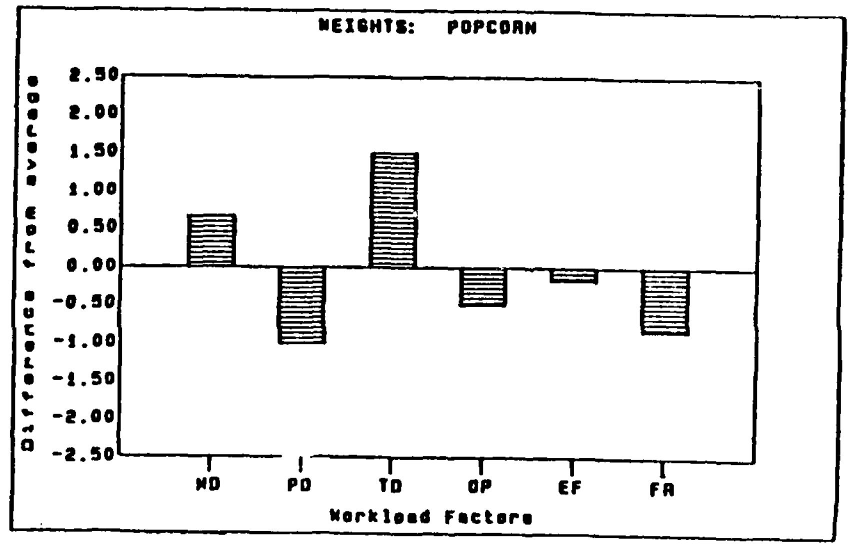




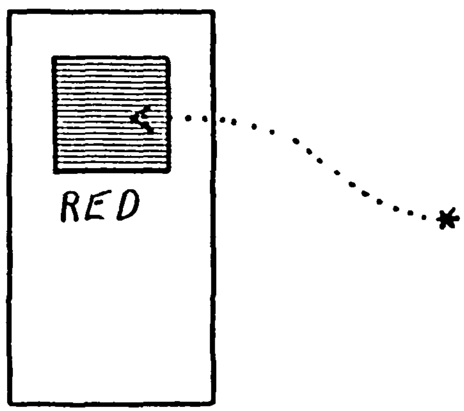


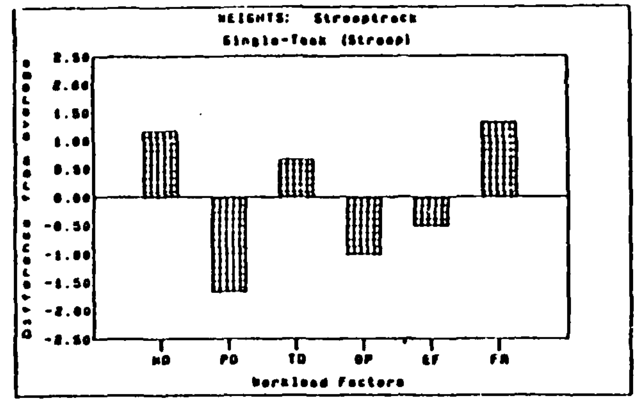
VALIDATION STUDY: POPCORN EXPERIMENT

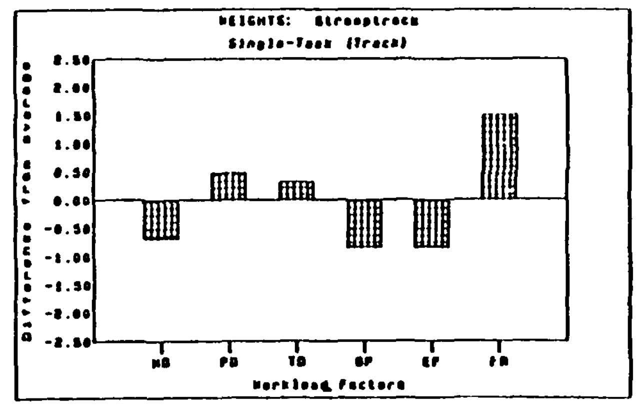


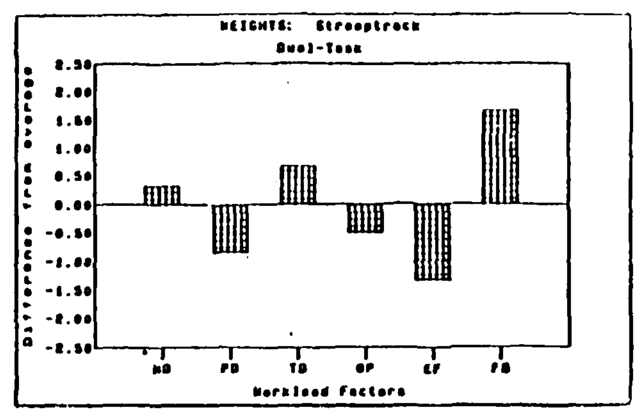


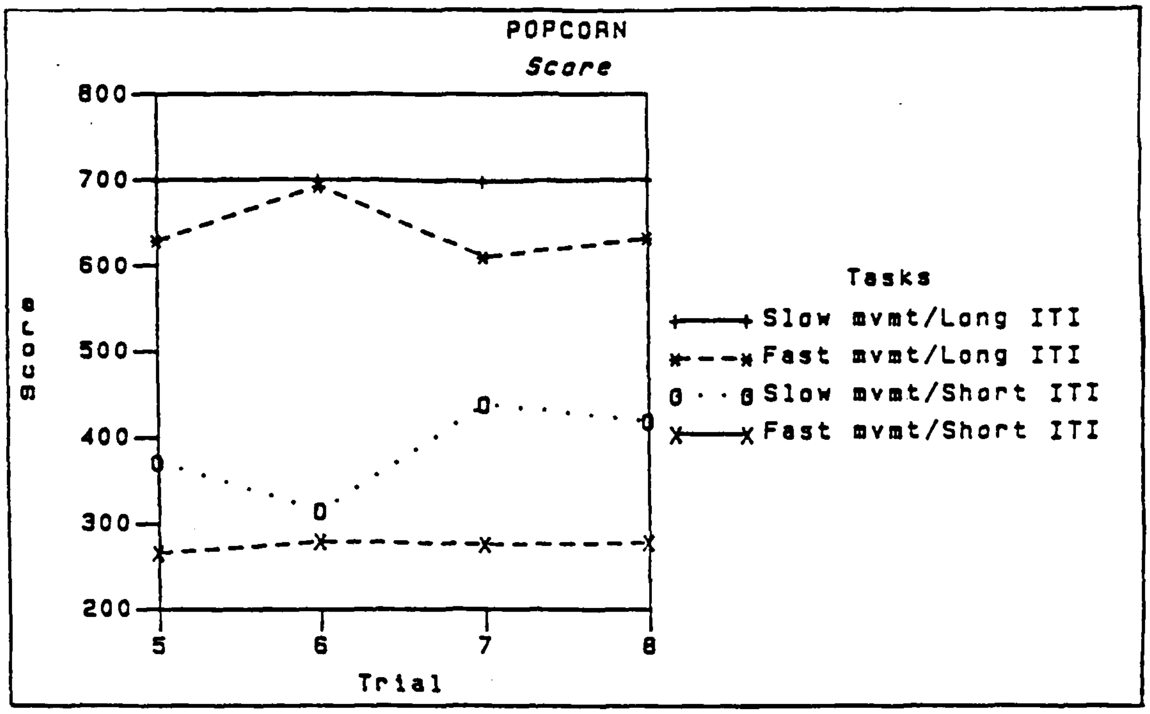
VALIDATION STUDY: STROOP TRACK

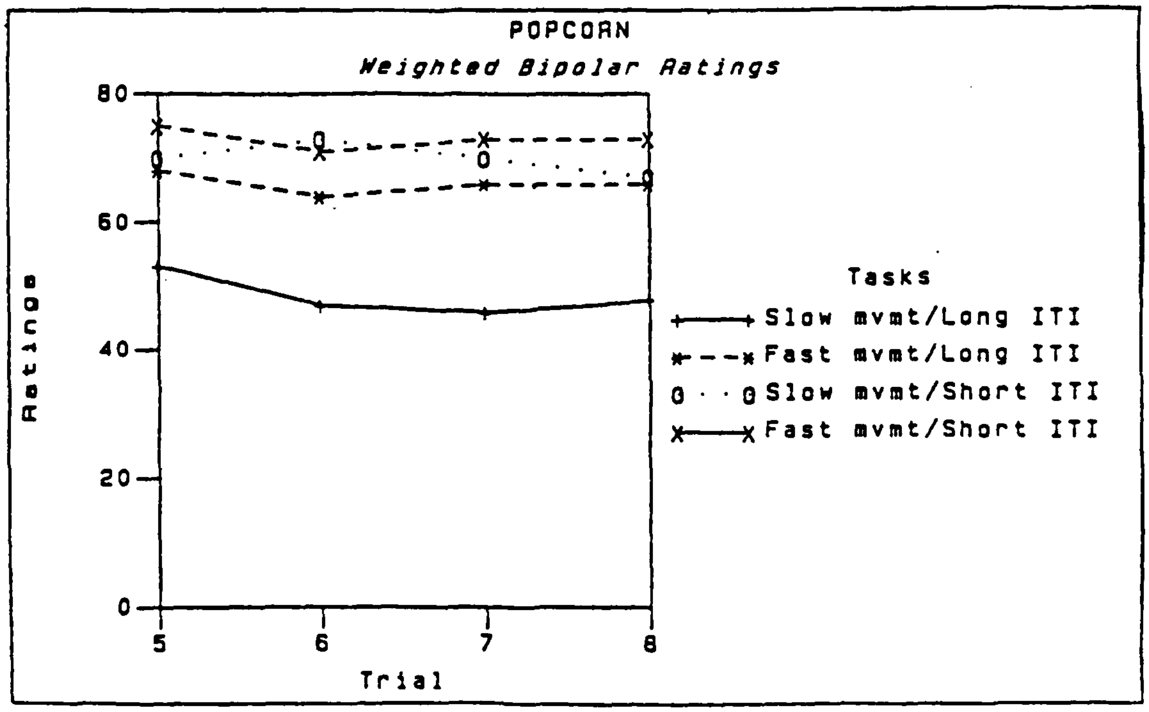




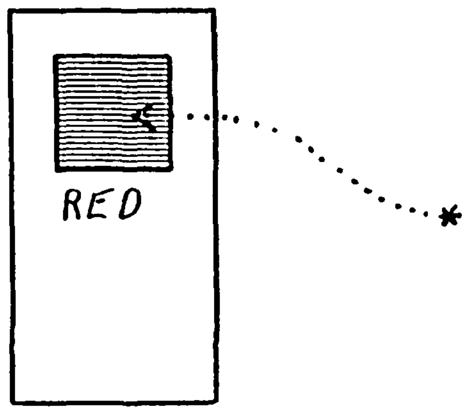


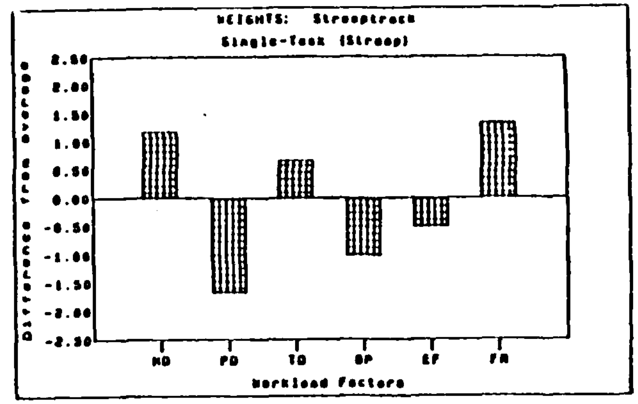


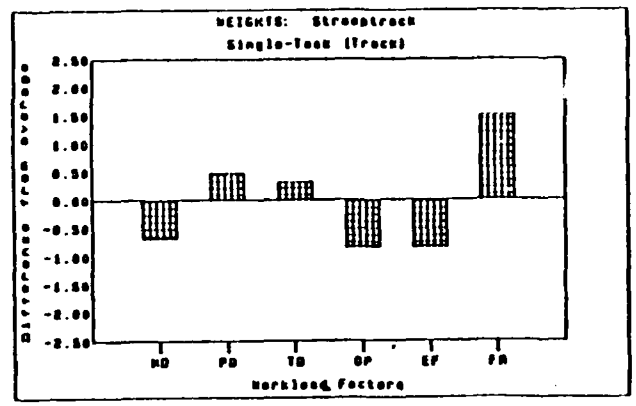

POPCORN
Score
800
700
600
500
400
300
200
Score
5
6
7
8
Trial
Tasks
Slow mvmt/Long ITI
Fast mvmt/Long ITI
Slow mvmt/Short ITI
Fast mvmt/Short ITI

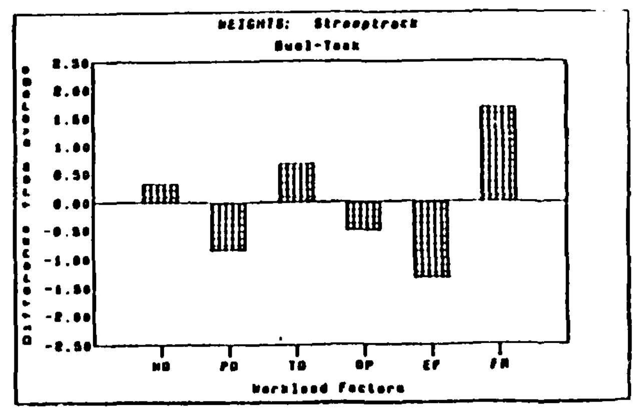

POPCORN
Weighted Bipolar Ratings
80
60
40
20
0
Ratings
5
6
7
8
Trial
Tasks
Slow mvmt/Long ITI
Fast mvmt/Long ITI
Slow mvmt/Short ITI
Fast mvmt/Short ITI

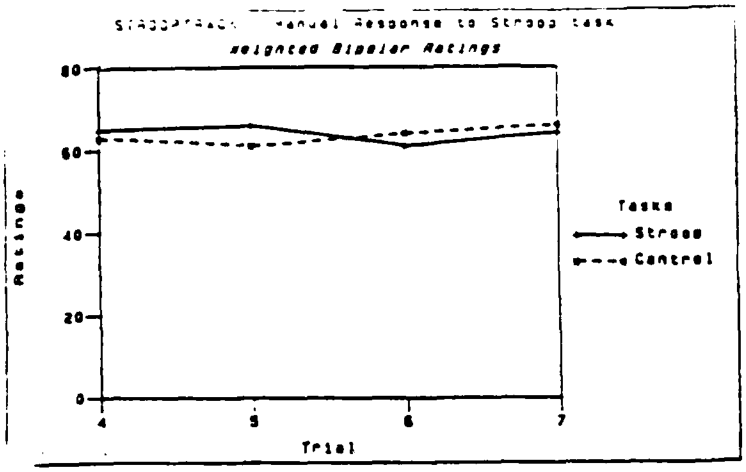
VALIDATION STUDY: STROOP TRACK

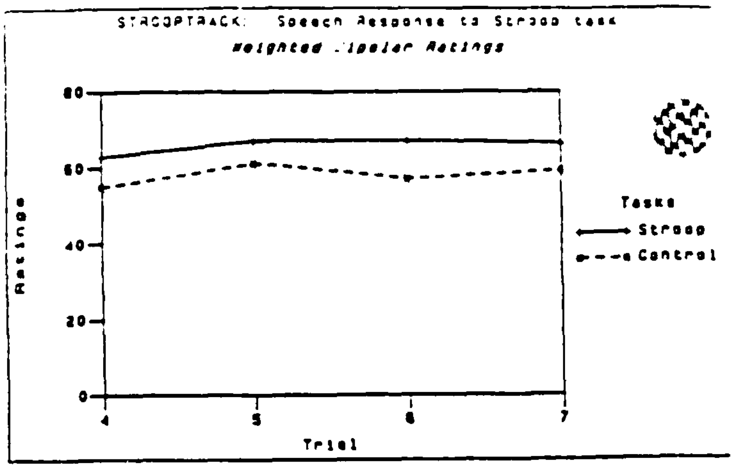


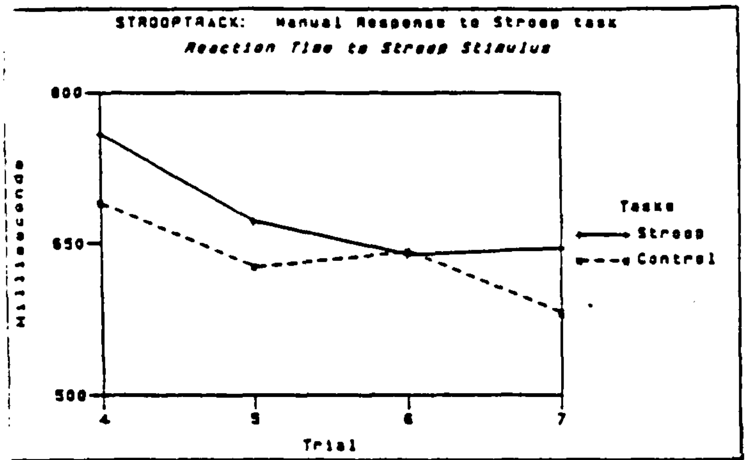


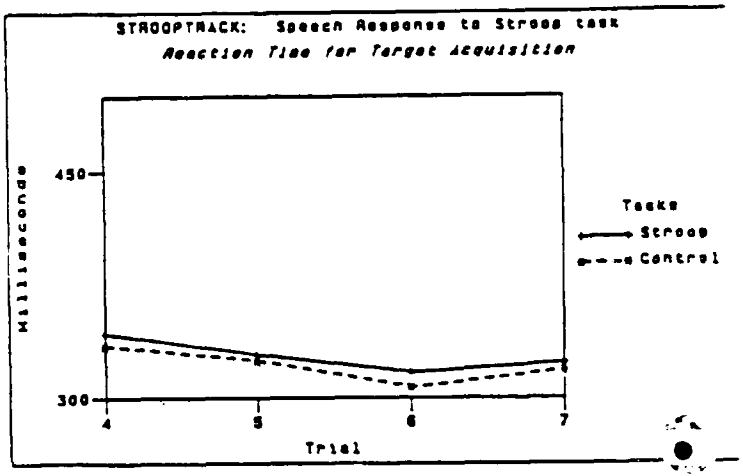


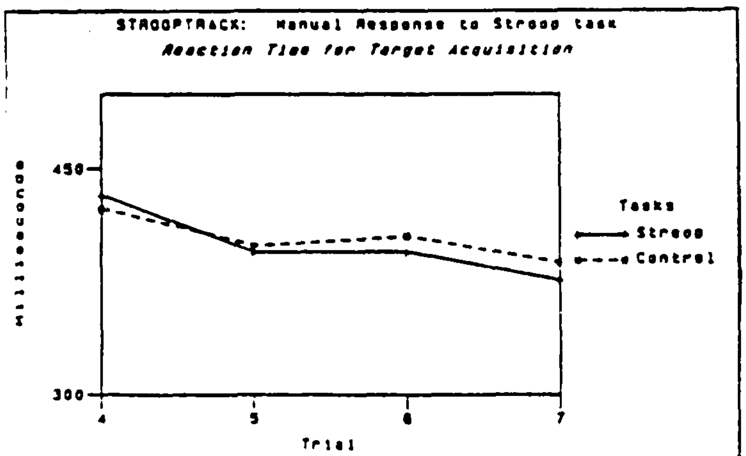


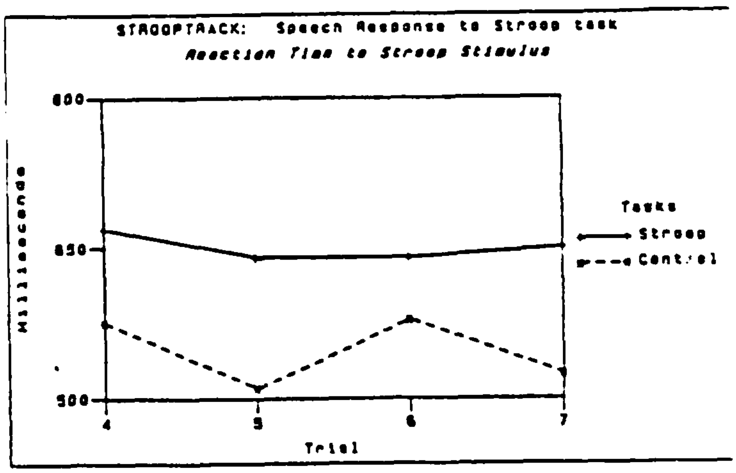

STROOPTRACK: Manual Response to Stroop task
Weighted Bipolar Ratings
Ratings
80
60
40
20
0
4
5
6
7
Trial
Tasks
Stroop
Control

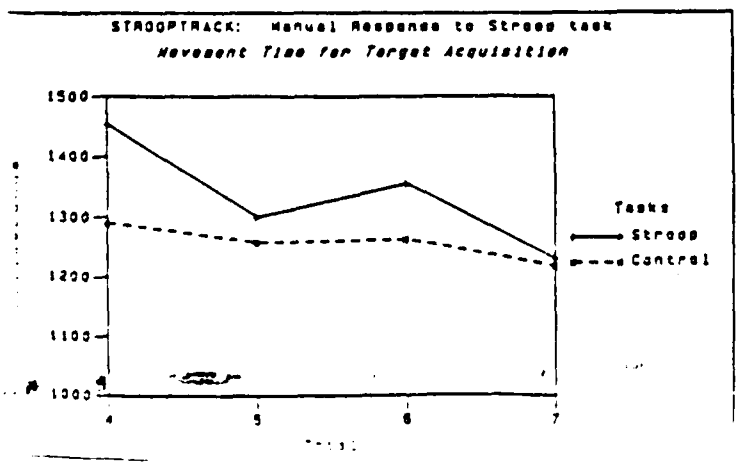

STROOPTRACK: Speech Response to Stroop task
Weighted Bipolar Ratings
Ratings
80
60
40
20
0
4
5
6
7
Trial
Tasks
Stroop
Control

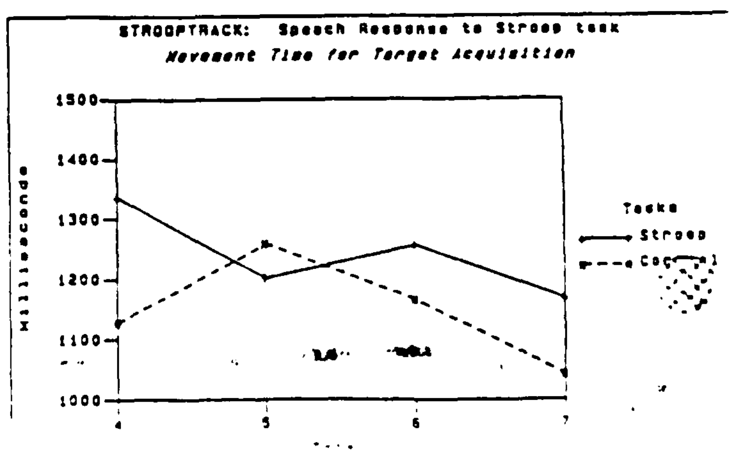

STROOPTRACK: Manual Response to Stroop task
Reaction Time to Stroop Stimulus
Milliseconds
800
650
500
4
5
6
7
Trial
Tasks
Stroop
Control


STROOPTRACK: Speech Response to Stroop task
Reaction Time for Target Acquisition
Milliseconds
450
300
4
5
6
7
Trial
Tasks
Stroop
Control


STROOPTRACK: Manual Response to Stroop task
Reaction Time for Target Acquisition
Milliseconds
450
300
4
5
6
7
Trial
Tasks
Stroop
Control


STROOPTRACK: Speech Response to Stroop task
Reaction Time to Stroop Stimulus
Milliseconds
800
650
500
4
5
6
7
Trial
Tasks
Stroop
Control


STROOPTRACK: Manual Response to Stroop task
Movement Time for Target Acquisition
1500
1400
1300
1200
1100
1000
4
5
6
7
Tasks
Stroop
Control


STROOPTRACK: Speech Response to Stroop task
Movement Time for Target Acquisition
Milliseconds
1500
1400
1300
1200
1100
1000
4
5
6
7
Tasks
Stroop

VALIDATION STUDY: COGNITIVE TASKS

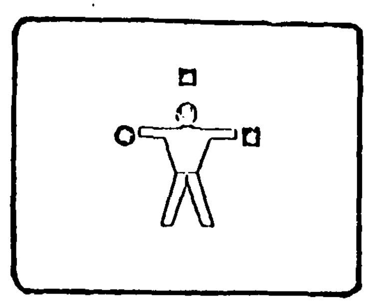
4	3	+	=	7 CORRECT ANSWER
9	4	-	=	5 CORRECT ANSWER
7	8	+	=	15 GREATER THAN NINE -10 SUBTRACT TEN 5 CORRECT ANSWER
2	9	-	=	-7 LESS THAN ZERO +10 ADD TEN 3 CORRECT ANSWER

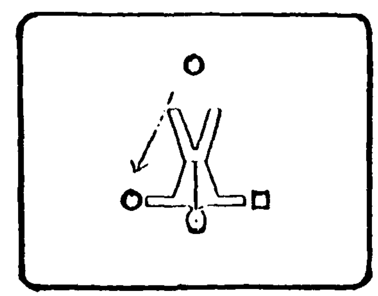
A STRING OF NINE NUMBERS IS PRESENTED. NO NUMBER APPEARS MORE THAN TWICE.

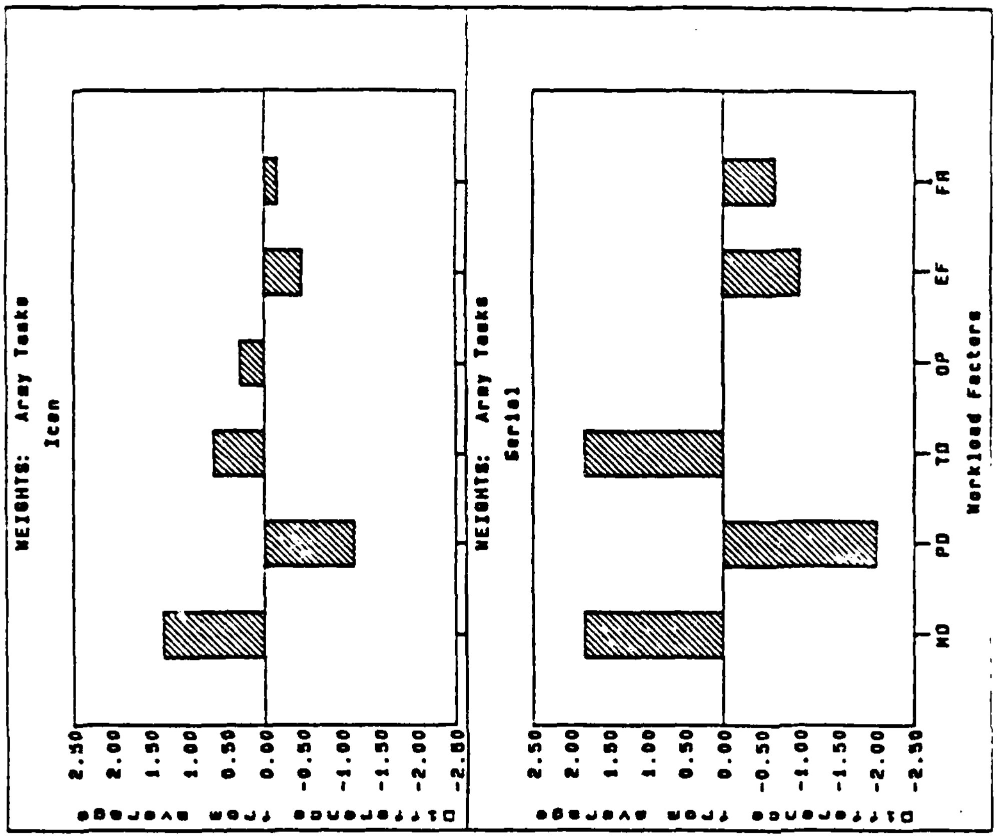
4 7 2 9 8 9 3 2 7

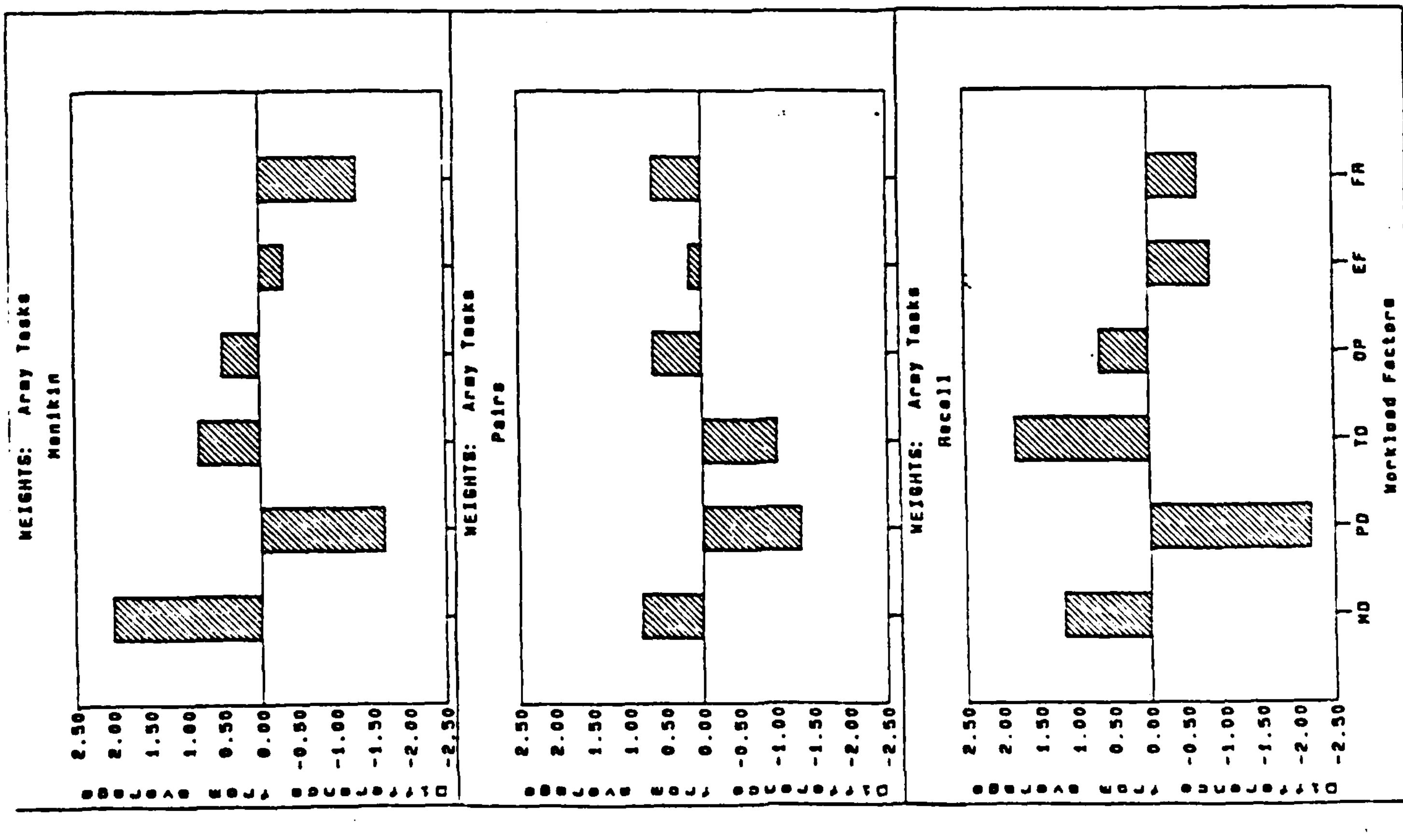
EIGHT OF THE NUMBERS PRESENTED IN SCRAMBLED ORDER. WHAT IS THE MISSING NINTH NUMBER?

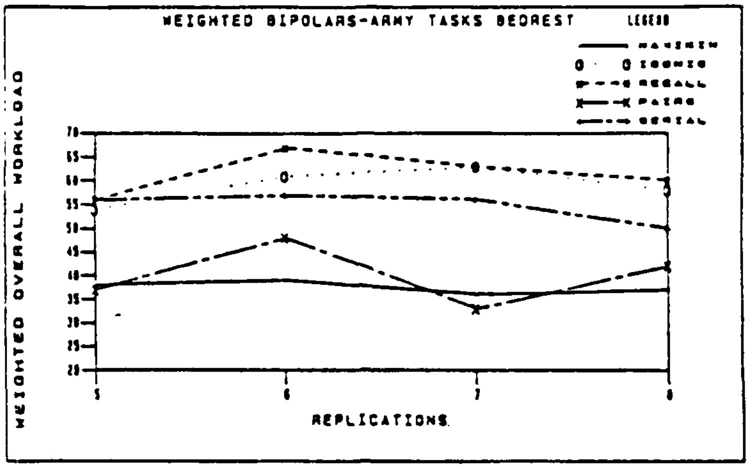
9 7 3 7 2 8 9 4

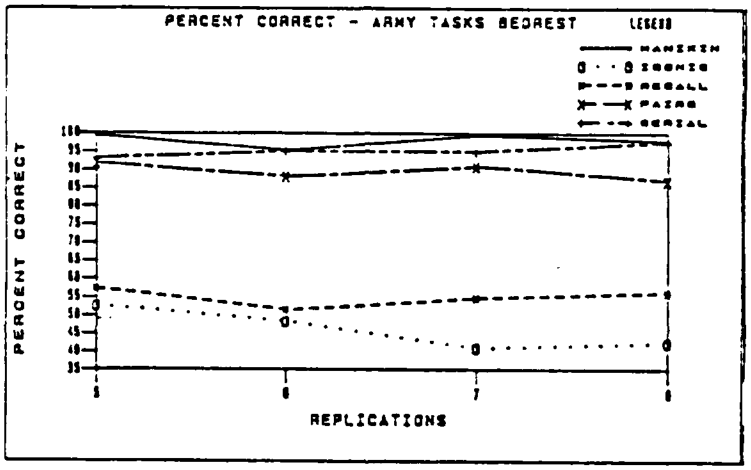


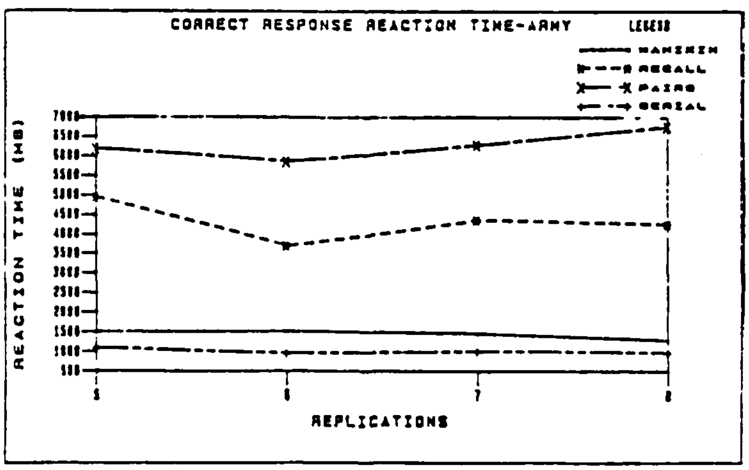



WEIGHTS: Army Tasks
Icon
WEIGHTS: Army Tasks
Serial
Difference from average
2.50
2.00
1.50
1.00
0.50
0.00
-0.50
-1.00
-1.50
-2.00
-2.50
ND
PD
TD
OP
EF
FR
Workload Factors

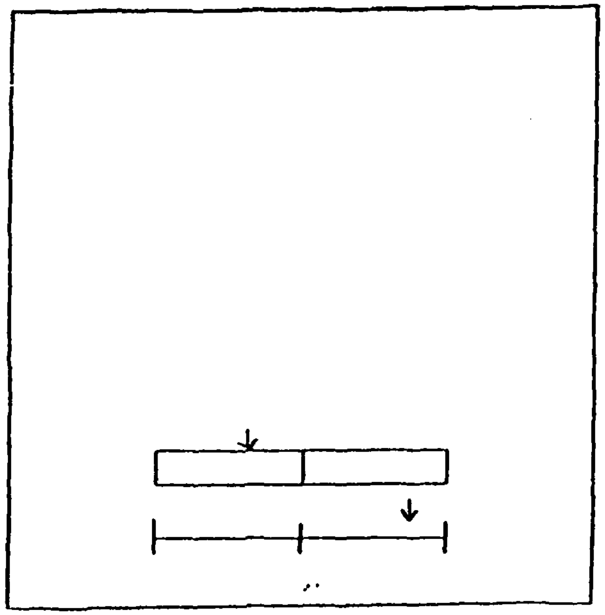

WEIGHTS: Army Tasks
Manikin
WEIGHTS: Army Tasks
Pairs
WEIGHTS: Army Tasks
Recall
Difference from average
2.50
2.00
1.50
1.00
0.50
0.00
-0.50
-1.00
-1.50
-2.00
-2.50
ND
PD
TD
OP
EF
FR
Workload Factors

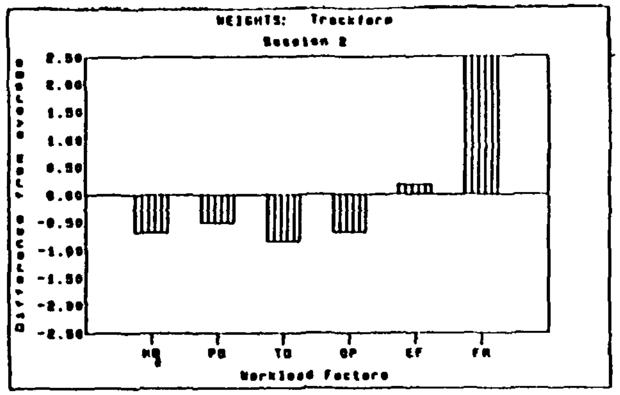

WEIGHTED BIPOLARS-ARMY TASKS BEDREST
MANIKIN
SIGNMIS
RECALL
PAIRS
SERIAL
WEIGHTED OVERALL WORKLOAD
REPLICATIONS

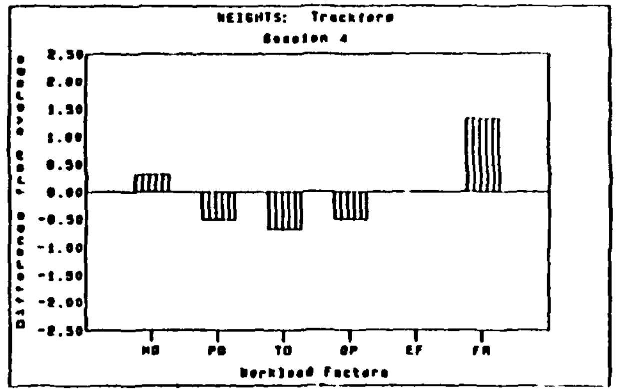

PERCENT CORRECT - ARMY TASKS BEDREST
MANIKIN
SIGNMIS
RECALL
PAIRS
SERIAL
PERCENT CORRECT
REPLICATIONS

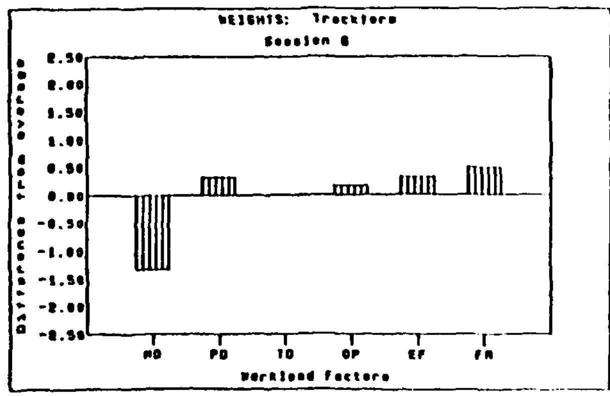

CORRECT RESPONSE REACTION TIME-ARMY
MANIKIN
RECALL
PAIRS
SERIAL
REACTION TIME (MS)
REPLICATIONS

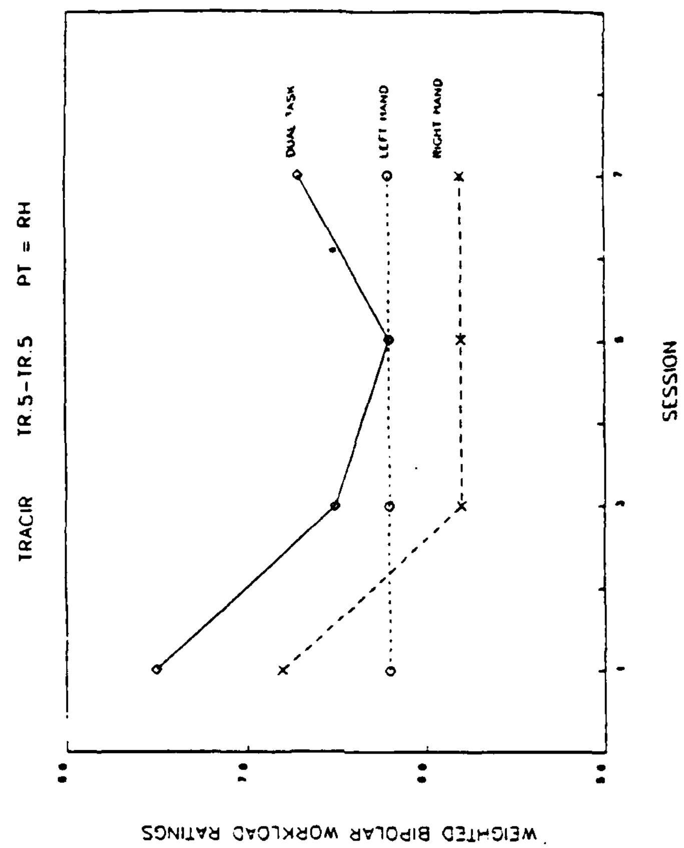
VALIDATION STUDY: TWO AXIS TRACKING TASK










TRACIR TR.5-TR.5 PT = RH
DUAL TASK
LEFT HAND
RIGHT HAND
SESSION
WEIGHTED BIPOLAR WORKLOAD RATINGS

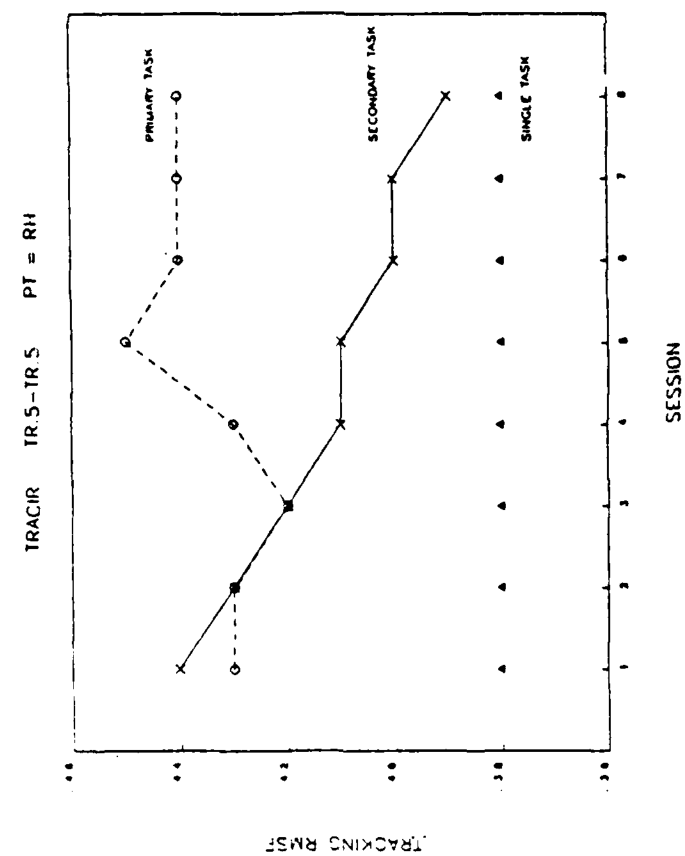

TRACIR TR.5-TR.5 PT = RH
PRIMARY TASK
SECONDARY TASK
SINGLE TASK
SESSION
TRACKING RMSE

Table 11: Validation Study

Correlations among bipolar ratings

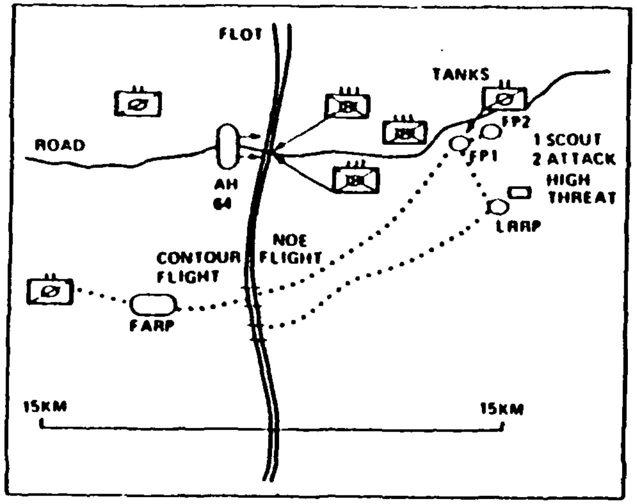
	MD	PD	TD	OP	EF	FR
PD	.57					
TD	.58	.50				
OP	.36	.27	.32			
EF	.76	.58	.66	.40		
FR	.54	.44	.52	.57	.69	
OW	.84	.70	.67	.46	.84	.70

Beta weights for the six rating subscales regressed on OW (*=p<.01)

	r2	MD	PD	TD	OP	EF	FR
SINGLE-COGNITIVE	.88	.43*	.15*	.04	.01	.33*	.13*
SINGLE-MANUAL	.78*	.38*	.39*	.11*	.12*	.21*	.00
DUAL-TASKS	.82	.41*	.19*	.02	.09*	.29*	.20*
FITTSBERG	.86	.32*	.24*	.17*	.09*	.16*	.19*
POPCORN	.90	.34*	.23*	.22*	.03	.19*	.10*
OVERALL	.86	.38*	.22*	.08	.05	.24*	.16*

HELICOPTER HUMAN FACTORS

SUPPORT FOR ARMY ADVANCED ROTORCRAFT TECHNOLOGY INTEGRATION (ARTI) PROGRAM



OBJECTIVE:

ASSESS THE WORKLOAD OF A SINGLE PILOT FLYING ADVANCED-TECHNOLOGY HELICOPTERS IN A SIMULATED COMBAT ENVIRONMENT

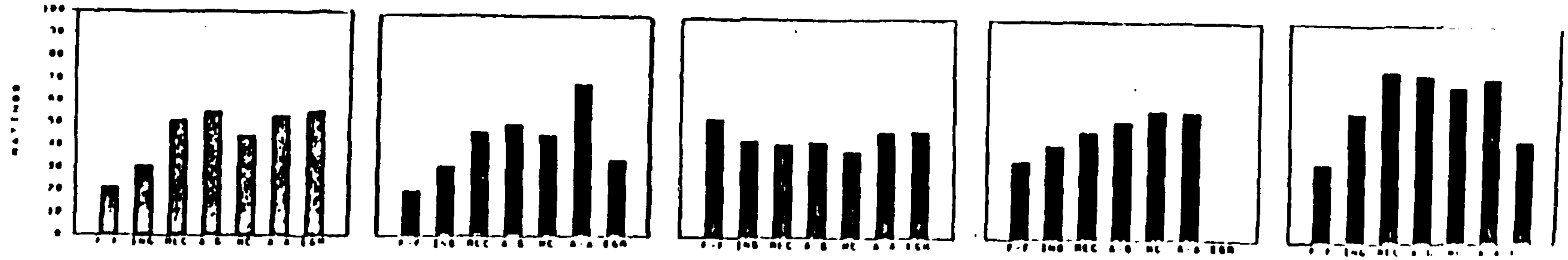
APPROACH:

— FLY SEVEN-SEGMENT GOVERNMENT COMPOSITE MISSION SCENARIO AT EACH OF FIVE ARTI CONTRACTOR SITES
— OBTAIN PILOT WORKLOAD RATINGS FOR EACH SEGMENT OF NORMAL AND DEGRADED FLIGHTS

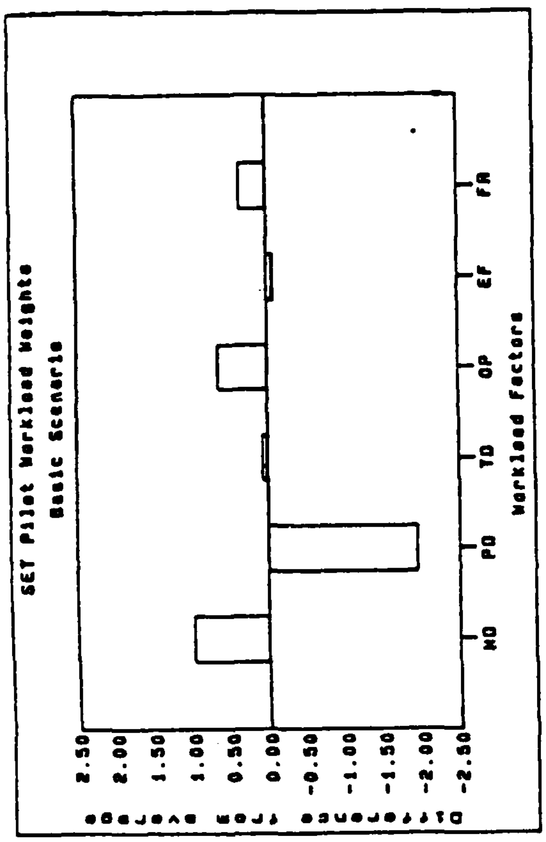
RESULTS:

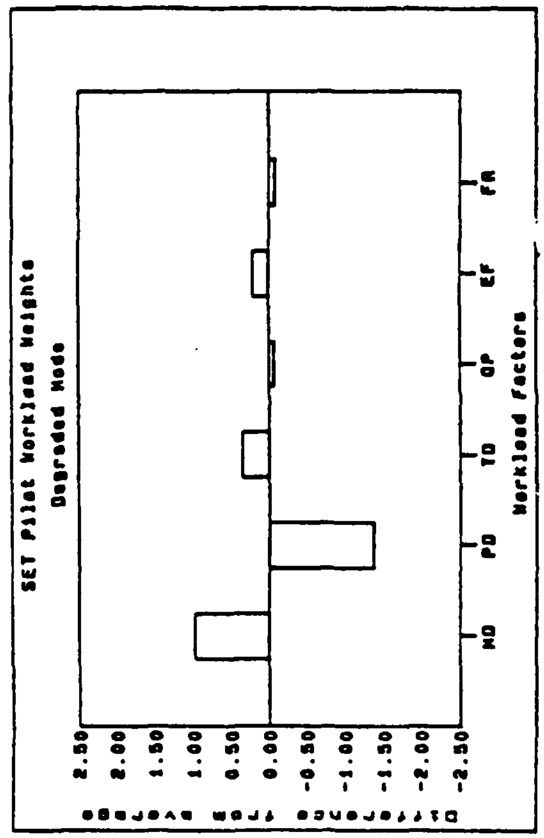
— AUTOMATION AND DISPLAY FEATURES REDUCED THE DEMANDS OF TRADITIONALLY HIGH SOURCES OF WORKLOAD, SUCH AS NAVIGATION
— SINGLE-PILOT WORKLOAD LEVELS REMAINED GENERALLY HIGH FOR AIR- TO-GROUND, RECONAISSANCE, AND AIR-TO-AIR SEGMENTS

WORKLOAD RATINGS FOR NORMAL MISSIONS FLOWN AT EACH SITE



FLIGHT SEGMENTS


SET Pilot Workload Weights
Basic Scenario
Workload Factors
Difference from average
2.50
2.00
1.50
1.00
0.50
0.00
-0.50
-1.00
-1.50
-2.00
-2.50
MD
PD
TD
OP
EF
FR

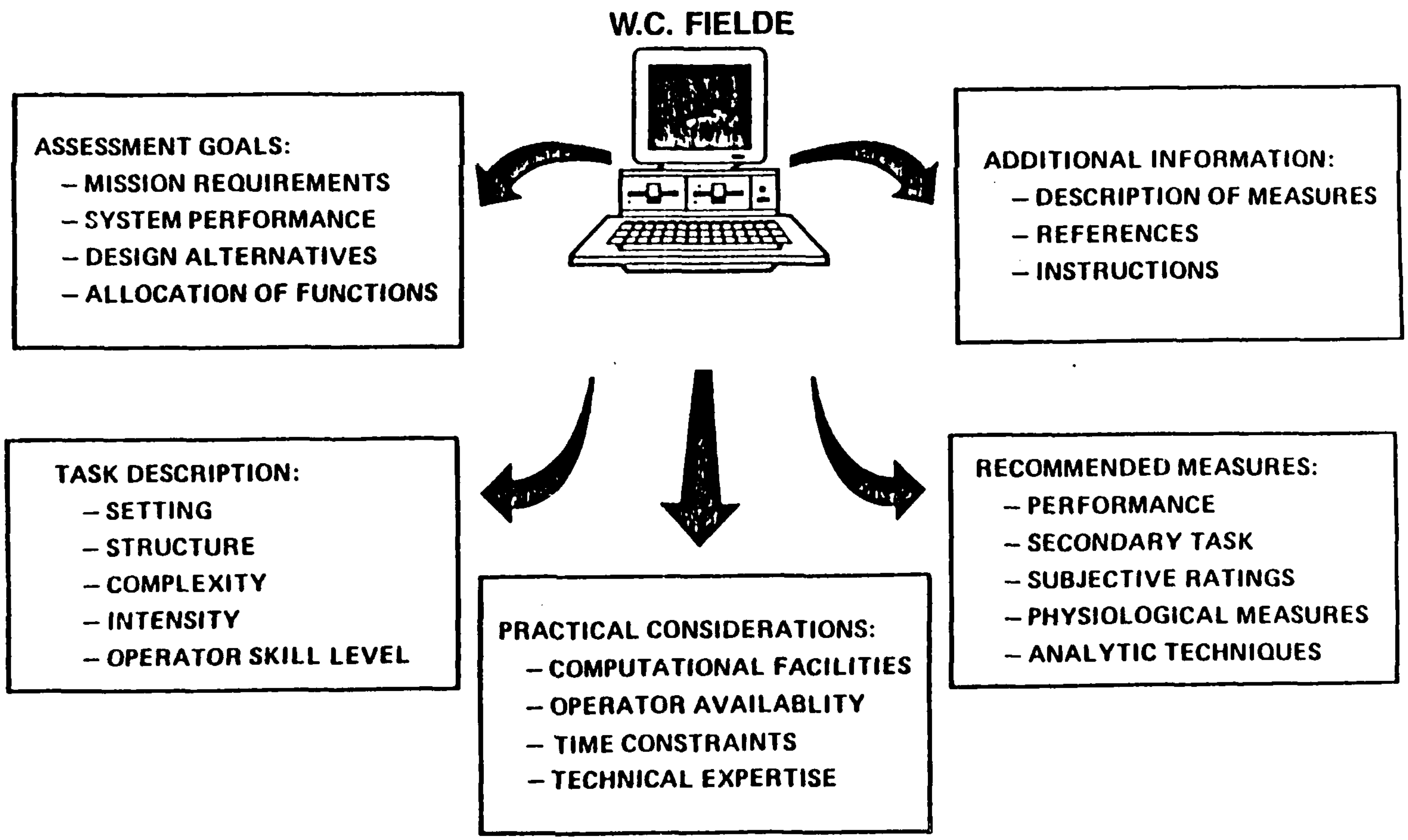

SET Pilot Workload Weights
Degraded Mode
Workload Factors
Difference from average
2.50
2.00
1.50
1.00
0.50
0.00
-0.50
-1.00
-1.50
-2.00
-2.50
MD
PD
TD
OP
EF
FR

WORKLOAD: SELECTION OF MEASURES

WHAT:	TASK ANALYSIS PERFORMANCE CRITERIA COMMUNICATIONS ANALYSIS ERRORS
WHEN:	TIME-LINE ANALYSIS RESPONSE TIME SUBJECTIVE RATINGS HEART RATE
HOW:	RESPONSE TIME ERRORS SUBJECTIVE RATINGS EVOKED CORTICAL POTENTIALS DIRECTION OF GAZE EYE BLINK RATE PERFORMANCE ON SECONDARY TASKS
WHERE:	PERFORMANCE DECREMENTS SUBJECTIVE RATINGS HEART RATE
WHO:	SUBJECTIVE RATINGS RELATIVE PERFORMANCE CREW INTERACTION ANALYSIS PHYSIOLOGICAL RESPONSES

W.C. FIELDE: Workload Consultant for FIELD Evaluation

AN INTERACTIVE EXPERT SYSTEM DESIGNED TO AID IN THE SELECTION AND APPLICATION OF WORKLOAD ASSESSMENT PROCEDURES



CONCLUSIONS

- WORKLOAD IS A MULTI-DIMENSIONAL CONSTRUCT
- SUBJECTIVE RATINGS DO NOT REPRESENT THE INHERENT PROPERTIES OF A TASK BUT EMERGE FROM THE INTERACTION BETWEEN AN OPERATOR, A TASK, AND THE ENVIRONMENT
- TASK-RELATED DIFFERENCES IN SOURCES OF WORKLOAD ARE BETTER PREDICTORS OF WORKLOAD EXPERIENCES THAN A PRIORI SUBJECTIVE BIASES
- A MULTI-DIMENSIONAL EVALUATION PROCEDURE THAT REFLECTS THE IMPORTANCE OF DIFFERENT FACTORS TO THE WORKLOAD OF A SPECIFIC TASK (THE WEIGHTS) AND THEIR MAGNITUDES (THE RATINGS) PROVIDE:
  - A SENSITIVE MEASURE OF OVERALL WORKLOAD
  - A REDUCTION IN BETWEEN-SUBJECT VARIABILITY
  - DIAGNOSTIC INFORMATION ABOUT THE WORKLOAD-STRUCTURE OF A TASK

SESSION 2. TASK ANALYSIS

Presenters

T. L. Ramirez and Lt. J. Masak, USAF: A Method for Determining Task Time Increase Caused by the Individual Protective Ensemble

J. Armstrong: Timebased Analysis of Significant Coordinated Operations (TASCO)

A Method for Determining Task Time Increase Caused by the Individual Protective Ensemble.

Tammy L. Ramirez, Robin L. Shew,
James E. Felt, and Michael E. Rayle
JAYCOR
3164 Presidential Dr. Fairborn, Ohio 45324

2Lt. Jerry Masak (USAF)
Harry G. Armstrong Aerospace Medical
Research Laboratory, Wright-Patterson
AFB, Dayton, Ohio 45433-6573

ABSTRACT

This report presents a methodology developed to measure task performance in a constrained environment [i.e., Individual Protective Equipment (IPE)]. The methodology has potential use in a number of human performance measurement areas involving increased time to complete tasks as a function of changes in the usual job environment. With moderate adjustments to an algorithm developed for this study, performance analysts can adapt this method to calculate time changes for other degraded job/task environments which exhibit similar characteristics, such as performance changes following the administration of pretreatment or antidotal drugs.

INTRODUCTION

BACKGROUND

The individual protective equipment (IPE) is an integral part of the protective posture assumed to facilitate continuance of operations in a chemically toxic environment, and is therefore a possible cause of job environment degradation which might produce substantial increases in task completion time. A review of the literature pertaining to human performance in a chemical environment revealed that the information is not sufficiently quantified to provide reliable expectations of task time increases. In the past, the procedure used to measure task time increase has been to: 1. observe tasks being performed by individuals wearing the IPE; 2. record the amount of time required to accomplish the task; 3. compare that time to a baseline measure which is typically the task time performance in a shirt-sleeve environment; 4. define the difference between the shirt-sleeve time and IPE time as the increase in time to perform. [Cox and Jeffers (1981); Hinch (1982)]. These studies along with other similar field studies (currently being performed by JAYCOR), are the beginning of a data base that may be used by simulation modelers.

IPE COMPONENTS

The IPE worn by Air Force ground support personnel consists of the following pieces of equipment: M17A1 protective mask, M6A2 protective hood, butyl rubber gloves with cotton liners, overgarment consisting of jacket and pants with charcoal interlining worn over fatigues, and overboots. Although the IPE provides protection for personnel in a toxic environment, the drawbacks of this protection are heat stress, encumbrance, degraded dexterity, restricted vision and lowered communication ability. As a result, the time required to complete many maintenance tasks is increased, and the capability to complete some tasks is seriously impaired.

SCOPE OF STUDY

This investigation was initiated specifically to develop a methodology for determining the task time increase for aircraft maintenance and munitions tasks as they apply to the Chemical Warfare Theater Simulation of Airbase Resources (CWTSAR) computer model.

The scope of this study included the following objectives:

- o Perform a review of the literature concerning measurement of human performance in a degraded environment.
- o Determine which human ability factors and task characteristics might best represent the components in a human performance model.
- o Develop a methodology for calculating the increase in time required to complete aircraft maintenance tasks while wearing IPE.
- o Determine task time multipliers that estimate the time increase to complete various aircraft maintenance tasks while wearing the IPE.
- o Implement a human performance data base for use by simulation modelers and job performance analysts.

A number of potentially interesting factors that were <u>not</u> considered in the method developed in this paper were:

- o Ambient temperature.
- o Work/rest cycles.
- o Safety factors in a degraded environment.
- o Skill levels of individuals performing the jobs.

These variables may eventually become part of the human performance data base, but are not considered in the scope of this study. A short discussion of the human performance model, the human abilities, and task framework selected to support the data base is presented in the following section.

HUMAN PERFORMANCE MODEL

The human performance model shown in Figure 1 graphically displays "someone, doing something, somewhere" [Bailey (1982)], but in this case the someone, something, and somewhere are constrained by the chemical warfare environment in which the tasks are being performed.

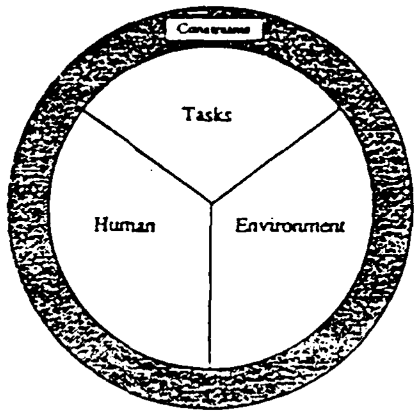


Figure 1 Human Performance Model

This illustration presents the reader with an example of the man-machine interface in a normal environmental setting (human, environment, tasks to be performed), which the data base represents as a baseline for task completion and human abilities. Bailey's model was then enclosed in constraints. The constraints are represented in the data base as difficulty factors. A likely consequence of increased difficulty is an increase in performance time.

HUMAN ABILITIES AND TASK DEFINITIONS

It is the premise of this study that human abilities can be rated according to the level of importance or "criticality" when performing tasks. The eight human ability categories and their corresponding subcategories were developed based on the task taxonomy and factor analytic work performed by Fleishman (1962, 1972, 1975, 1978, and 1982). These abilities are assumed to adequately characterize the demands of tasks that are critical to a military maintenance environment. The eight abilities (see table 1) describe basic functions needed to begin, continue, and complete a task.

As discussed by Fleishman (1967), the term "ability" refers to a general trait which has been inferred from consistent responses on certain tasks. The human performance ability categories in Table 1 form the basis for the subjective rating measures (of criticality and difficulty) used to determine task time multipliers (TTM).

A task, as defined in this study, is one of the major subdivisions of a job performed by an individual. A job usually consists of between two and seven tasks. The following are some characteristics of a task:

- It is one of the main functions required for the personnel to perform in order to complete the job successfully.
- It comprises a grouping of closely related subtasks.
- Task requirements often are the basis for initial assignment to a job, and for determining the qualifications required to perform the job.

TABLE ONE. HUMAN PERFORMANCE ABILITY CATEGORIES

Auditory Detection	Dexterity	Psychological Effects
localization	fine motor manipulation	stress
sensitivity	fine motor response	tension
response rate	fine motor strength	depression
speech interference		anxiety
intensity		confusion
		motivation

Cognitive Effects	Vision
short term	acuity
long term	accommodation
retention	distance
storage	perception
concentration	color discrimination
attention	peripheral vision

Communication	Physiological Conditions	Physical Coordination
understand speech	fatigue	motor response
response process	stamina	general mobility
	adaptation	strength

A practical method for estimating increase in task performance time in a degraded environment can be based on subjective estimates of the criticality of human abilities (Gebhardt, et. al., 1981) to task performance. It is also assumed that increases in difficulty (as compared to a normal environment) due to wearing IPE can be subjectively estimated by job incumbents. The present method obtains subjective estimates of criticality and increased difficulty (using a five point scale for both measures) for each of the human ability categories shown in Table 1. These estimates are then combined (as described below) in order to obtain an overall Task Time Multiplier.

METHODOLOGY

A flow diagram (Figure 2) displays the process employed for development of the data base. This diagram also provides the reader with an overview of the steps which lead to the design of the task time multiplier algorithm.

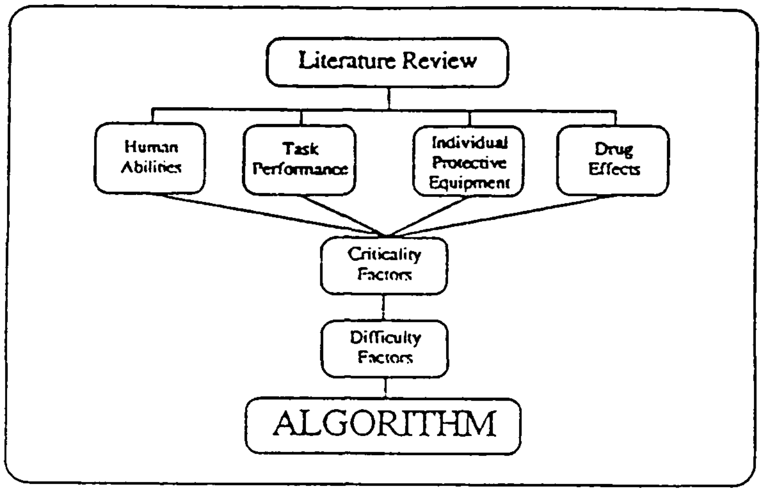


Figure Two Flow Diagram

The eight human abilities tentatively identified as representative of the components required for task performance included in this data base are listed in Table 2 with their respective definitions. It may appear that parts of definitions overlap or fit into other categories, but until more research dictates further refinement, these definitions will be used.

TABLE 2 Ability Definitions

1. Auditory detection	The sensitivity to sounds (other than verbal communication), and interpretation of sounds.
2. Cognitive effects	Ability to concentrate and attend to a task. Includes the short-term and long-term memory, retention, and recall of information.
3. Communication	Ability to understand and respond to speech for relating ideas necessary for task completion.
4. Dexterity	Fine motor response, fine motor manipulation (ability to work by touch when unable to see an object), fine motor strength.
5. Physical coordination	The ability to climb, drive, walk, use general mobility, and muscular strength.
6. Physiological conditions	Tactile pressure, fatigue, personal needs, and stamina experienced in performance of tasks.
7. Psychological effects	Levels of stress, tension, depression, anxiety, confusion, and motivation.
8. Vision	Use of acuity, accommodation, depth perception, adaptation, color discrimination, and peripheral vision to perform a task.

HUMAN PERFORMANCE MATRIX

A representation of the matrix used to display the various factors is shown in Figure 3. The matrix displays, as an example, vehicle mechanic tasks. The contrived task percentage, human performance abilities, criticality factors, and difficulty factors have been arbitrarily assigned. These values would, in theory, yield the TTM for each task the vehicle mechanic performs. A discussion of each section of the matrix beginning with the task percentage is presented below.

Speciality Code:
Job Name:

Task Performance	PATCH HOLES IN AIRCRAFT		FABRICATE AIRCRAFT PARTS													
Task Percentage																
Task Time																
Criticality and Difficulty Factors	C/F	D/F	C/F	D/F	C/F	D/F	C/F	D/F	C/F	D/F	C/F	D/F	C/F	D/F	C/F	D/F
Human Abilities: Auditory Detection																
Cognitive Effects																
Communication																
Dexterity																
Physical Coordination																
Physiological Conditions																
Psychological Effects																
Vision																
Product Sum																
Criticality Sum																
Performance Number																
Task Time Multiplier																
Total Increased Time to Perform																
Conversion Scale	Performance number TTM	5 2.2	4.5 2.05	4 1.9	3.5 1.75	3 1.6	2.5 1.45	2 1.3	1.5 1.15	1 1						

Figure 3 TTM Matrix

TASK PERCENTAGE

The tasks that comprise the particular job being measured are presented across the top of Figure 3. Unlike the abilities discussed earlier, tasks will change according to the job being measured. Task specifications can be found in references such as Air Force Regulation (AFR) 39-1, "Airman Classification." Percentage of task performance is defined as the percent of the total job that an individual spends performing the task. For example, the tasks of a vehicle mechanic are:

- o Tuning engines
- o Repairing exhaust systems
- o Repairing suspension systems
- o Repairing electrical systems
- o Adjusting and repairing brakes

The mechanic may spend 30 percent of the job time tuning engines and 5 percent repairing exhaust systems. It is assumed that most job classifications can be viewed in this manner. The performance analyst may evaluate any task which is performed. Those tasks possessing high criticality and high difficulty factors, may need equipment or task redesign to enhance performance in a degraded environment. Conversely, a task which is performed infrequently, and has low criticality and difficulty factors, is one that may not need further investigation.

CRITICALITY FACTOR

The criticality factor is defined as the level of importance of a human ability to the performance of a task, development of this factor consists of establishment of the relative importance of each of the eight abilities to completion of a task. A questionnaire being developed by JAYCOR will collect the criticality data. This questionnaire will be distributed to personnel performing the tasks and they will subjectively rate the importance of each factor to task completion. The questionnaire uses a five point scale, one equal to not important and five equal to very important.

DIFFICULTY FACTOR

The level of difficulty is the amount of constraint placed on an individual by changes in the environment. The difficulty factor is a function not only of the task, but also of the particular source of task degradation (IPE, pretreatment, etc.). The procedure for determining the difficulty factor is also a subjective rating (one equal to not difficult and five equal to very difficult) obtained from the personnel performing the task.

The criticality factors are rated in order to establish a weighting factor for each of the human abilities. The difficulty ratings are obtained following task completion while wearing the IPE. It is not necessary to have the same individual rate the criticality and the difficulty factors for the same task. Both the criticality and difficulty factors are measures of central tendency ($\bar{x}$) for a number of individuals and therefore are averages of a set of data.

TASK TIME MULTIPLIER

The TTM for each task is arrived at by employing the following algorithm:

- For each of the eight human abilities, the criticality factor is multiplied by the difficulty factor.
- These products are summed over the eight human abilities.
- The sum of the products is then divided by the sum of the eight criticality factors to give the performance number.
- The performance number is then converted to the TTM, by linearly rescaling its range of possible values, 1 to 5, to

the TTM range, 1 to 2.2 [i.e., TTM = 1 + .97 (P.N. - 1)] using the conversion scale discussed below.

CONVERSION SCALE DEVELOPMENT

The conversion scale, used in the above algorithm, provides the transition from the Performance Number to the TTM. This scale, as displayed at the bottom of Figure 3 allows for the linear rescaling of the performance number to the task time multiplier. This scale has been developed from operational field data pertaining to munitions and maintenance tasks. Eighty-five task times were compiled from reports of operational chemical defense training exercises. Of these 85 tasks measured when wearing the IPE, only one task time exceeded twice the normal completion time. Based on these data, it was empirically determined to set the highest value of the conversion scale at 2.2. Therefore, if tuning an engine takes 80 minutes and is increased by a TTM of 1.97, then the increased task time would be 157.5 minutes. This scale is in the process of being refined to associate the performance number to a TTM more accurately. As additional data become available, the conversion scale will be adjusted to reflect any changes.

CONCLUSIONS AND RECOMMENDATIONS

CONCLUSIONS

The conclusions of this study are:

- o Human abilities and task characteristics for a human performance model can be roughly defined.
- o A methodology for calculating the increased time for tasks in a degraded environment has been developed.

It is further concluded that:

- o Data collection for various aircraft maintenance tasks is required in order to develop a more comprehensive data base and to test various simulation models.
- o Implementation of a human performance data base for use with simulation models is of value and enhances the utility of the models.

The human performance data base developed through this study provides the following:

- o A method of calculating increased time to perform the task.

- o Identification of tasks within the job which may exceed reasonable time to complete.
- o Identification of human abilities that are being stressed and that may be high contributors to task time increase.
- o Tasks that may need further evaluation or the redesign of equipment as a method of decrease task time.

RECOMMENDATIONS

Although progress has been made, there is still considerable validity and reliability testing required. As data are collected and added to the data base the strength of this methodology will be tested. The following recommendations are presented as conditions for implementation and testing of the human performance data base:

- o Continue to solicit opportunities for operational field data collection.
- o Implement software to perform data runs for preliminary validation of the human performance data base.
- o Continue refinement of the conversion scale for TTM reliability.

REFERENCES

AFR 39-1, Airman Classification, Change 8, 1982 (January): Manpower Personnel Command, Randolph Air Force Base, San Antonio, Texas.

Bailey. R.W., 1982:Human Performance Engineering: A Guide for System Designers., Bell Telephone Laboratories, Prentice-Hall Publishing, Inc., Englewood Cliffs, New Jersey.

Cox, T.J. and A.R. Jeffers, 1981: Ground Crew Chemical Defense Equipment Performance Task Time Degradation Test (U), ASD-TR-81-5003, Wright-Patterson AFB, Ohio, AD B 057 406.

Fleishman, E.A., 1982 (July): "Systems for describing human tasks," American Psychologist, Vol. 37, pp. 821-834.

Fleishman, E.A. and J.C. Hogan, 1978 (June): "Taxonomic method for assessing the physical requirements of jobs: The physical abilities approach," ARRO Tech Report 3012/R 78-6. Advanced Research Resources Organization, Washington, D.C.

Fleishman E.A., 1975: "Toward a Taxonomy of Human Performance," American Psychologist, Vol. 30. pp. 1127-1149.

Fleishman, E.A., 1972: "On the Relation Between Abilities, Learning, and Human Performance," American Psychologist, Vol. 27, pp. 1017-1032.

Fleishman, E.A., 1967: "Individual Differences in Motor Learning" In R.M. Gagne (ed.), Learning and Individual Differences, Charles E. Merrill Publishing, Columbus, Ohio.

Fleishman, E.A. and G.D. Ellison, 1962: "A Factor Analysis of Fine Manipulative Performance,"Journal of Applied Psychology, Vol. 46, pp. 96-105.

Gebhardt, D.L., M.C. Jennings, E.A. Fleishman, 1981 (February): Factors Affecting the Reliability of Physical Ability and Effort Ratings of Navy Tasks , AD A098 333.

Hinch, T.J., 1982: Performance of Airbase Personnel in Chemical Protective Clothing, USAF Assistant Chief of Staff, Studies and Analyses, Washington, D.C.

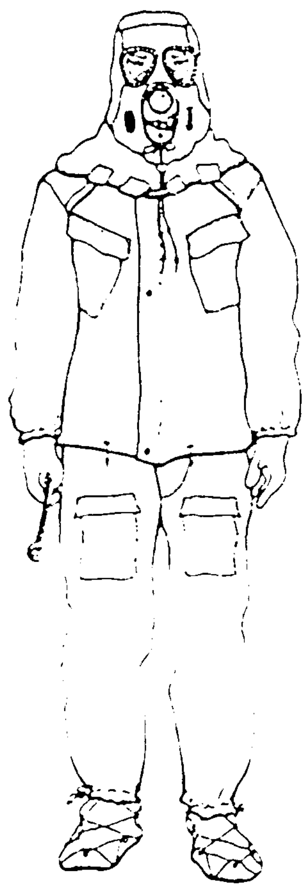
JAYCOR

Human Performance Database

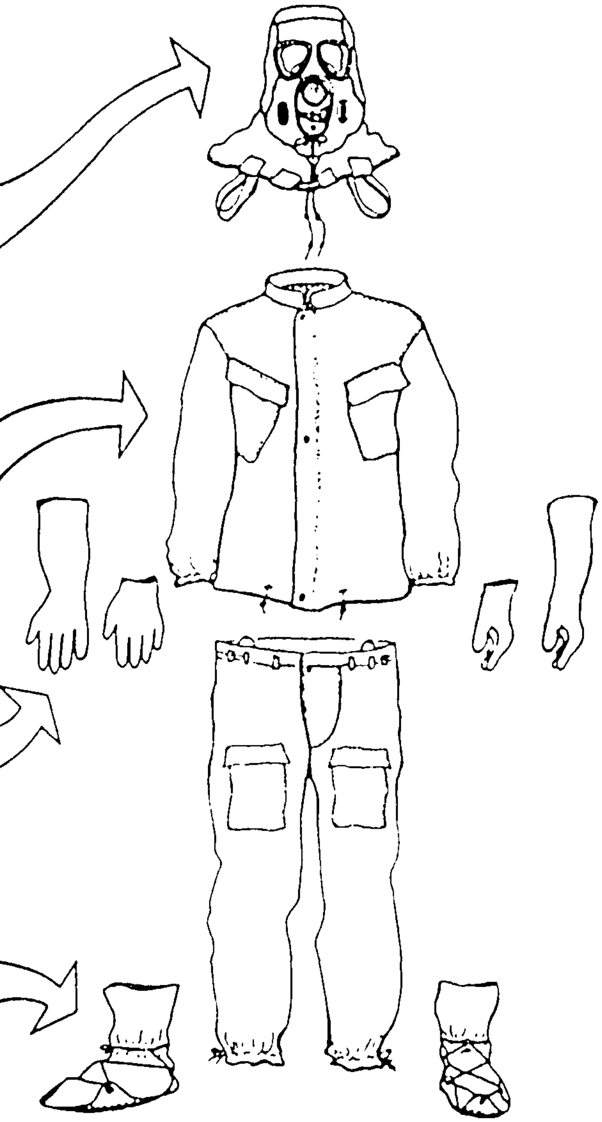
- Task Time Multiplier Development
- Human Performance Model
- Job Analysis
- Human Abilities

JAYCOR

IPE Degradation



EQUIPMENT	EFFECT ON PERSONNEL
M6A2 Chemical Protective Hood and M17A1 Protective Mask	Communication problems Encumbrance Discomfort Vision difficulties Thermal stress
Chemical Protective Suit	Encumbrance Dexterity loss Thermal stress
Chemical Protective Glove Set	Encumbrance Dexterity loss
Chemical Protective Footwear	Encumbrance Dexterity loss



JAYCOR

- Task Time Multiplier Development

  Human Performance Model

  - Task
  - Human
  - Environment
  - Constraints

  Job Analysis

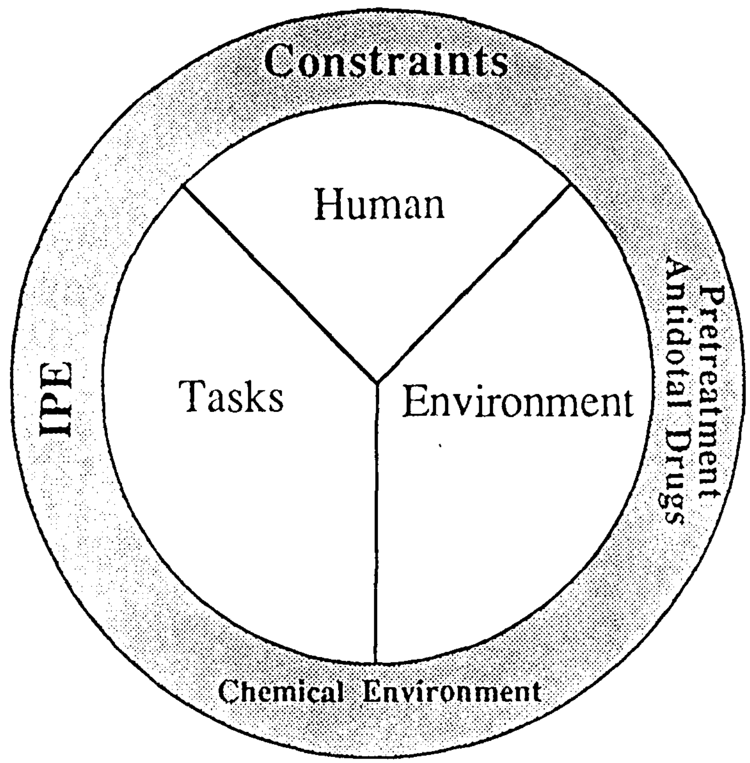
  - Task Performance
  - Criticality of Performance
  - Difficulty of Performance

  Human Abilities

  - Performance Measure
  - Select Abilities of Interest
  - Subcategories of Abilities

JAYCOR

Human Performance Model



JAYCOR

JOB ANALYSIS DEFINITIONS

- Task = a subdivision of a job performed by personnel
- Criticality = the level of importance attributed to a human ability in performance of a task
- Difficulty = the amount of burden placed on an individual by changes or CONSTRAINTS in the environment

- Performance Measures
  - Full Scale Mission Simulation (Grodsky, 1967)
    - Very Good , But Costly
  - Synthetic-work Methodology (Alluisi, 1967)
    - Excellent for Training Studies
  - Factor-Analysis Techniques (Fleishman, 1967)
    - Best for Our Purpose (Human Performance and Constraints)

JAYCOR

Abilities Of Interest

Eight Human Performance Ability Categories And Selected Subcategories

1 Auditory Detection

- localization
- sensitivity
- response rate
- speech interference
- intensity

2 Communication

- understand speech
- response process

3 Cognition

- short term
- long term
- retention
- storage
- concentration
- attention

4 Dexterity

- fine motor manipulation
- fine motor response
- fine motor strength

5 Physiological Conditions

- fatigue
- stamina
- adaptation

6 Physical Coordination

- motor response
- general mobility
- strength

7 Psychological Conditions

- stress
- tension
- depression
- anxiety
- confusion
- motivation

8 Vision

- acuity
- accommodation
- distance
- perception
- color discrimination
- peripheral vision

JAYCOR

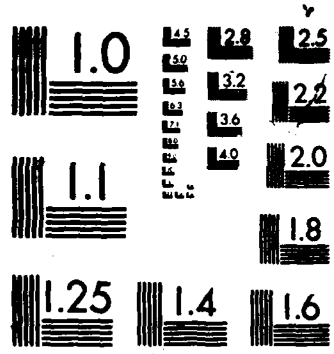
Questionnaire

- Criticality of Ability and Subcategory of Ability
- Difficulty of Ability and Each Subcategory
- In Laymen's Terms.

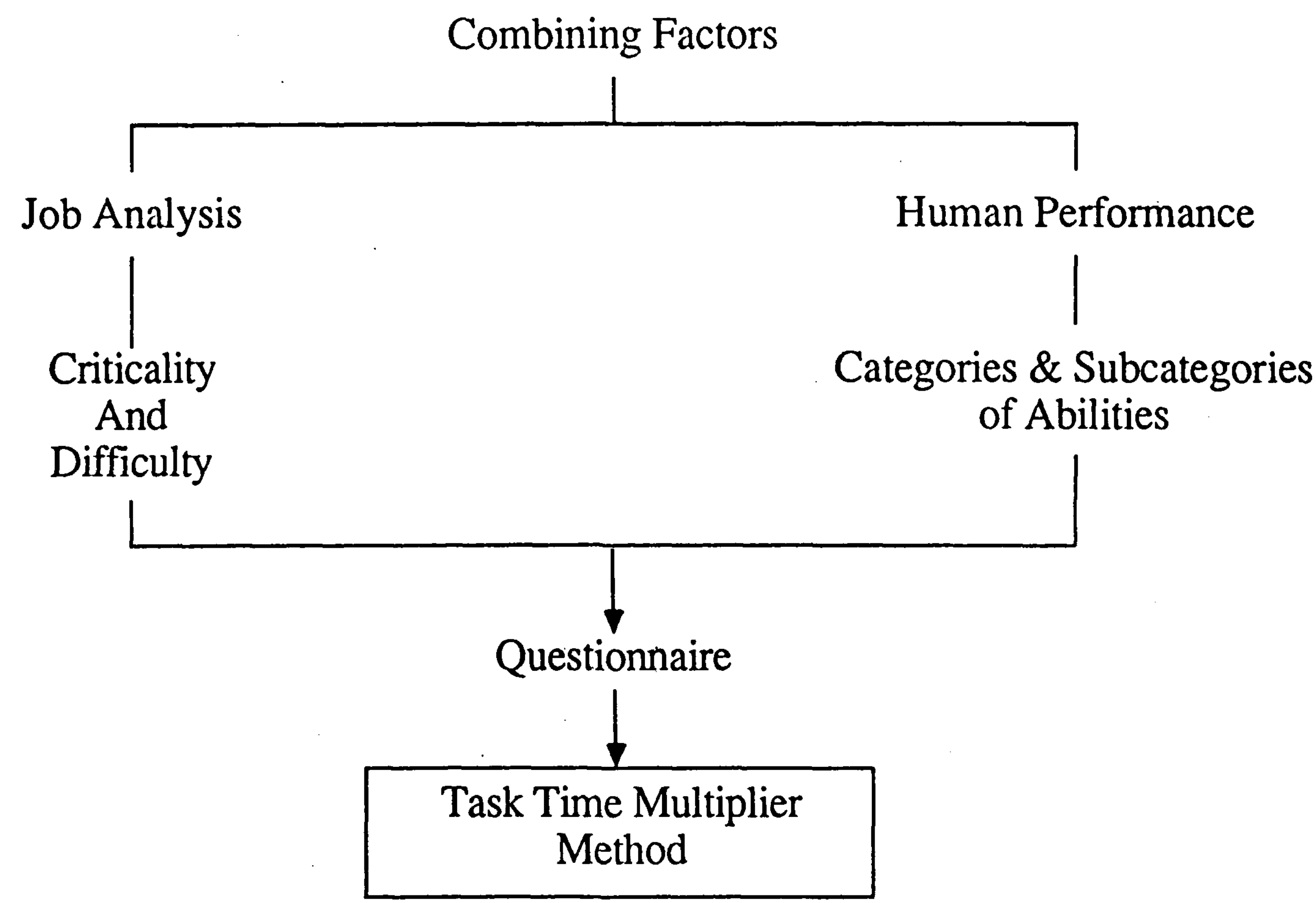
AD-A185 650 PROCEEDINGS OF THE DOD WORKLOAD ASSESSMENT WORKSHOP ON WORKLOAD ASSESSMEN..(U) NAVAL UNDERWATER SYSTEMS CENTER NEWPORT RI H M FIEDLER 15 SEP 87 NUSC-TD-6608

3/4

UNCLASSIFIED F/G 23/2 NL


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0

JAYCOR


Combining Factors
Job Analysis
Human Performance
Criticality
And
Difficulty
Categories & Subcategories
of Abilities
Questionnaire
Task Time Multiplier
Method

Example TTM

Speciality Code: 47252
Job Name: Vehicle Mechanic

	Task Performance	TUNING ENGINES		REPAIRING EXHAUST SYSTEMS		REPAIRING SUSPENSION SYSTEMS		REPAIRING ELECTRICAL SYSTEMS		REPAIRING & ADJUSTING BRAKES							
	Task Percentage	30%		5%		5%		40%		*							
	Task Time	80 min.		150 min.		60 min.		240 min.		30 min.							
	Criticality and Difficulty Factors	C/F	D/F	C/F	D/F	C/F	D/F	C/F	D/F	C/F	D/F	C/F	D/F	C/F	D/F	C/F	D/F
Human Abilities	Auditory Detection	4.0	4.0														
	Cognitive Effects	3.0	3.0														
	Communication	2.0	4.0														
	Dexterity	4.0	5.0														
	Physical Coordination	2.0	3.0														
	Physiological Conditions	5.0	5.0														
	Psychological Effects	2.0	3.0														
	Vision	4.0	5.0														
	Product Sum	110															
	Criticality Sum	26															
	Performance Number	4.23															
	Task Time Multiplier	1.97															
	Total Increased Time to Perform	157.5 min.															

Conversion Scale										
Performance Number	5	4.5	4	3.5	3	2.5	2	1.5	1
TTM	2.2	2.05	1.9	1.75	1.6	1.45	1.3	1.15	1

JAYCOR

TTM Factors Compared to Actual Time Increases

	TTM Factors	Actual Increase Factor
Munitions	1.46	1.47
Maintenance	1.55	1.5
Civil Engineering	1.63	1.7

Conclusions

- Human Abilities For Air Force Tasks Have Been Roughly Defined.
- Methodology For Calculating Increased Performance Time Has Been Developed.
- Method For Studying Subcategories of Human Abilities Has Been Developed.
- This Procedure May Prove Useful For Other Tasks Performed In Changed Environments.

RETURN TO:

HARRY G. ARMSTRONG AEROSPACE MEDICAL RESEARCH LABORATORY
SPECIAL PROJECTS OFFICE
AEROSPACE MEDICAL DIVISION
AIR FORCE SYSTEMS COMMAND (AAMRL/HET)
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433-6573

ATTENTION: 2 LT. MASAK

Autovon: 785-7583

Commercial: (513) 255-7583

BIOGRAPHICAL DATA SHEET

Please fill out the biographical data sheet as accurately and completely as possible. All of your comments will be kept confidential.

DATE: ____________

I. MILITARY INFORMATION:

1. AFSC No.: ________ 2. Job Title ________________
3. Rank: ____________ 4. Duty Phone: ______________
5. Length of Time in Active Military Duty: ________
   _______ years and _______ months
6. Number of Chemical Warfare Exercises You Have Participated in
   ____
7. Number of Chemical Warfare Exercises You Have Participated in during the Past Year ________
8. Number of Months You Have Been Stationed at this Base
   ____

II. PHYSICAL INFORMATION:

1. Age: ________ (years)
2. Height: ________ (feet), ________ (inches)
3. Weight: ________ (pounds)
4. Sex: ________ Male ________ Female

III. Work Conditions

During combat conditions where would you perform your job?

a. Inside
b. Outside
c. Both

Explain ____

____

HUMAN PERFORMANCE QUESTIONNAIRE

PURPOSE

The purpose of this questionnaire is to gather information about various aspects of your AFSC. The data collected will be used to determine how your job performance might be affected by such factors as wearing the chemical defense ensemble, and other changes in your work environment.

INSTRUCTIONS

1. To complete the questionnaire, circle the appropriate number corresponding to your answer. For example:

   How important is clear vision to the performance of your job?

   1 Not important
   2 Somewhat important
   3 Important
   4 Very important
   5 Most important

2. Again, all responses will be confidential.

3. If you have any questions, please ask the interviewer.

1. How important is clear vision to the performance of your job?
   1. Not important.
   2. Somewhat important.
   3. Important.
   4. Very important.
   5. Most important.

2. How important is the ability to quickly focus your eyes (up close and then far or far then up close) to the performance of your job?
   1. Not important
   2. Somewhat important.
   3. Important.
   4. Very important.
   5. Most important.

3. How important is the ability to see objects at a distance?
   1. Not important.
   2. Somewhat important.
   3. Important.
   4. Very important.
   5. Most important.

4. How important to your job is depth perception (three-dimensional)?
   1. Not important.
   2. Somewhat important.
   3. Important.
   4. Very important.
   5. Most important.

5. How important is it for your eyes to adjust to varying amounts of light (brightness/darkness)?
   1. Not important.
   2. Somewhat important.
   3. Important.
   4. Very important.
   5. Most important.

6. How important to the performance of your job is the ability to distinguish different colors?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

7. How important to the performance of your job is the ability to watch or monitor instruments or gages?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

8. Overall, how important is vision to the performance of your job?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

9. How important to your job is the ability to hear emergency sounds? (List specific emergency sounds.)
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

10. How important to your job is the ability to determine volume levels (sound)?

    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

11. How important to your performance are workplace noise levels?

    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

12. How important is sensitive hearing to the performance of your job (e.g., audio equipment)?

    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

13. How important is the ability to hear equipment sounds (e.g., troubleshooting equipment sounds)?

    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

14. How important to your job is the ability to clearly hear others speaking?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

15. How important is it for you to react to sounds?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

16. Overall, how important is hearing to the performance of your job?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

17. How important is a sense of balance to the performance of your job?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

18. How important is being fully rested to the performance of your job?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

19. How important to your job is the ability to do physical work for long periods of time?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

20. Overall, how important is your physical condition to the performance of your job?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

21. How important to your job is climbing?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

22. How important to your job is driving?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

23. How important to your job is walking?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

24. How important is general mobility to the performance of your job?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

25. How important is muscular strength to your job?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

26. Overall, how important are physical capabilities to the performance of your job (e.g., driving, walking)?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

NOTE: The following five questions measure influence instead of importance.

27. How much influence does stress or tension have on your job performance?
    1. Not influenced.
    2. Somewhat influenced.
    3. Influenced.
    4. Very influenced.
    5. Most influenced.

28. How much influence does depression (e.g., sadness, despair) have on the performance of your job?
    1. Not influenced.
    2. Somewhat influenced.
    3. Influenced.
    4. Very influenced.
    5. Most influenced.

29. How much influence does anxiety have on the performance of your job (e.g., worried, concerned)?
    1. Not influenced.
    2. Somewhat influenced.
    3. Influenced.
    4. Very influenced.
    5. Most influenced.

30. How much influence does feeling confused have on the performance of your job (e.g., not understanding directions for given job)?
    1. Not influenced.
    2. Somewhat influenced.
    3. Influenced.
    4. Very influenced.
    5. Most influenced.

31. How much influence does motivation have on the performance of your job?
    1. Not influenced.
    2. Somewhat influenced.
    3. Influenced.
    4. Very influenced.
    5. Most influenced.

32. How important is the ability to recall procedures (Tech Orders)?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

33. How important is the ability to quickly recall correct procedures to the performance of your job (e.g., emergency, safety procedures)?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

34. How important to your job is the ability to concentrate under stress?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

35. How important is the ability to focus attention to the job ~~at hand~~?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

36. How important is the ability to quickly perform logical processes (e.g., math, technical procedures)?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

37. <u>Overall</u>, how important is memory to the performance of your job?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

38. How important to your job performance is the ability to understand what someone is saying?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

39. How important to your job is the ability to speak clearly?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

40. How important to your job is the ability to write legibly?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

41. <u>Overall</u>, how important to your job is the ability to communicate?
    1. Not important.
    2. Somewhat important.
    3. Important.
    4. Very important.
    5. Most important.

42. How important is dexterity to the performance of your job?

1. Not important.
2. Somewhat important.
3. Important.
4. Very important.
5. Most important.

43. Using the following scale, please rate each of the items as it pertains to your job:

1. Not important.
2. Somewhat important.
3. Important.
4. Very important.
5. Most important.

______ Making Keystrokes (typewriter, computer, calculator)
______ Pushing Buttons, knobs
______ Flicking Switches
______ Using Clamps
______ Manipulating Levers
______ Using Screwdrivers, tools
______ Turning Dials
______ Using Keys
______ Using a Firm Grip
______ Making Steady Wrist Movements
______ Aligning Equipment
______ Tracking Objects Across a Screen
______ Inserting Objects or Instruments into Small Places
______ Threading Objects (Stripped Bolts, Needles)

INSTRUCTIONS

The following questions are designed to gather information about the difficulties you may have experienced while performing your job in IPE. The data will be used to help to eliminate some of these difficulties. To complete the questionnaire, circle the appropriate number corresponding to your answer.

How difficult is it to see clearly while wearing IPE?

1. Not difficult.
2. Somewhat difficult.
3. Difficult.
4. Very difficult.
5. Most difficult.

If you have any questions, please ask the interviewer.

1. How difficult is it to see clearly while wearing IPE?
   1. Not difficult.
   2. Somewhat difficult.
   3. Difficult.
   4. Very difficult.
   5. Most difficult.

2. How difficult is it to quickly focus your eyes (up close and then far or far then up close) while wearing IPE?
   1. Not difficult
   2. Somewhat difficult.
   3. Difficult.
   4. Very difficult.
   5. Most difficult.

3. How difficult is it to see objects at a distance while wearing IPE?
   1. Not difficult.
   2. Somewhat difficult.
   3. Difficult.
   4. Very difficult.
   5. Most difficult.

4. How difficult is it to use depth perception (three-dimensional vision) while wearing IPE?
   1. Not difficult.
   2. Somewhat difficult.
   3. Difficult.
   4. Very difficult.
   5. Most difficult.

5. How difficult is it for your eyes to adjust to varying amounts of light (brightness/darkness) while wearing IPE?
   1. Not difficult.
   2. Somewhat difficult.
   3. Difficult.
   4. Very difficult.
   5. Most difficult.

6. How difficult is it to distinguish different colors while wearing IPE?
   1. Not difficult.
   2. Somewhat difficult.
   3. Difficult.
   4. Very difficult.
   5. Most difficult.

7. How difficult is it to watch or monitor instruments or gages while wearing IPE?
   1. Not difficult.
   2. Somewhat difficult.
   3. Difficult.
   4. Very difficult.
   5. Most difficult.

8. Overall, how difficult is it to see while wearing IPE?
   1. Not difficult.
   2. Somewhat difficult.
   3. Difficult.
   4. Very difficult.
   5. Most difficult.

9. How difficult is it to hear emergency sounds while wearing IPE? (List specific emergency sounds.)
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

10. How difficult is it to determine volume levels (sound) while wearing IPE?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

11. How much influence do work place noise levels have on the performance of your job while wearing IPE?
    1. No influence.
    2. Some influence.
    3. Moderate influence.
    4. High influence.
    5. Excessive influence.

12. How difficult is it to clearly distinguish different sounds while wearing IPE (e.g., audio equipment)?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

13. How difficult is it to hear equipment sounds (e.g., troubleshooting equipment sounds) while wearing IPE?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

14. How difficult is it to clearly hear others speaking while wearing IPE?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

15. How difficult is it for you to react to sounds while wearing IPE?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

16. Overall, how difficult is it to hear while wearing IPE?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

17. While wearing IPE, how difficult is it to maintain a sense of balance?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

18. How difficult is it to perform your job in the IPE while feeling tired?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

19. How difficult is it to do physical work for long periods of time while wearing IPE?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

20. <u>Overall</u>, how difficult would it be to perform your job in the IPE if in poor physical condition?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

21. How difficult is climbing while wearing IPE?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

22. How difficult is driving while wearing IPE?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

18. How difficult is it to perform your job in the IPE while feeling tired?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

19. How difficult is it to do physical work for long periods of time while wearing IPE?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

20. Overall, how difficult would it be to perform your job in the IPE if in poor physical condition?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

21. How difficult is climbing while wearing IPE?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

22. How difficult is driving while wearing IPE?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

23. How difficult is walking while wearing IPE?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

24. How difficult is general movement while wearing IPE?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

25. How difficult is it to maintain muscular strength while wearing IPE?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

26. Overall, how difficult is it to maintain good physical proficiency while wearing IPE?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

NOTE: The following five questions are on aspects other than difficulty.

27. How much stress or tension is caused by the IPE?
    1. No stress.
    2. Some stress.
    3. Moderate stress.
    4. Extreme stress.
    5. Severe stress.

28. How depressed (e.g., sadness, despair) do you feel while wearing IPE?
    1. No depression.
    2. Some depression.
    3. Moderate depression.
    4. Extreme depression.
    5. Severe depression.

29. How much anxiety do you feel while wearing IPE (e.g., worried or concerned)?
    1. No anxiety.
    2. Some anxiety.
    3. Moderate anxiety.
    4. Extreme anxiety.
    5. Severe anxiety.

30. How much confusion do you feel while wearing IPE (e.g., not understanding directions, for given job)?
    1. No confusion.
    2. Some confusion.
    3. Moderate confusion.
    4. Extreme confusion.
    5. Severe confusion.

27. How much stress or tension is caused by the IPE?
    1. No stress.
    2. Some stress.
    3. Moderate stress.
    4. Extreme stress.
    5. Severe stress.

28. How depressed (e.g., sadness, despair) do you feel while wearing IPE?
    1. No depression.
    2. Some depression.
    3. Moderate depression.
    4. Extreme depression.
    5. Severe depression.

29. How much anxiety do you feel while wearing IPE (e.g., worried or concerned)?
    1. No anxiety.
    2. Some anxiety.
    3. Moderate anxiety.
    4. Extreme anxiety.
    5. Severe anxiety.

30. How much confusion do you feel while wearing IPE (e.g., not understanding directions, for given job)?
    1. No confusion.
    2. Some confusion.
    3. Moderate confusion.
    4. Extreme confusion.
    5. Severe confusion.

31. How is your motivation to perform your job well affected by wearing IPE?
    1. No motivation.
    2. Some motivation.
    3. Moderate motivation.
    4. Strong motivation.
    5. High motivation.

32. How difficult is it to recall procedures (Tech Orders) while wearing IPE?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

33. How difficult is it to quickly recall correct procedures while wearing IPE (e.g., emergency safety procedures)?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

34. How difficult is it to concentrate under stress while wearing IPE?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

35. How difficult is it to focus attention to the job at hand while wearing IPE?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

36. How difficult is it to quickly perform logical processes while wearing IPE (e.g., math, technical procedures)?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

37. Overall, how difficult is it to remember things while wearing IPE?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

38. How difficult is it to understand what someone is saying while wearing IPE?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

39. How difficult is it to speak clearly while wearing IPE?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

40. How difficult is it to write legibly while wearing IPE?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

41. Overall, how difficult is it to communicate while wearing IPE?
    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

42. How difficult is it to perform tasks which require dexterity while wearing IPE?

    1. Not difficult.
    2. Somewhat difficult.
    3. Difficult.
    4. Very difficult.
    5. Most difficult.

43. Using the following scale, please determine how difficult it is to perform each task while wearing IPE:

1. Not difficult.
2. Somewhat difficult.
3. Difficult.
4. Very difficult.
5. Most difficult.

______ Making Keystrokes (typewriter, computer, calculator)
______ Pushing Buttons, knobs
______ Flicking Switches
______ Using Clamps
______ Manipulating Levers
______ Using Screwdrivers, tools
______ Turning Dials
______ Using Keys
______ Using a Firm Grip
______ Making Steady Wrist Movements
______ Aligning Equipment
______ Tracking Objects Across a Screen
______ Inserting Objects or Instruments into Small Places
______ Threading Objects (Needles, Stripped Bolts)

227

Timebased Analysis of Significant Coordinated Operations
(TASCO)
by J. Armstrong

TIMEBASED ANALYSIS OF SIGNIFICANT COORDINATED OPERATIONS

A COMPUTERIZED TASK ANALYSIS DATA BASE

by J. Armstrong

1.0 INTRODUCTION

A systems approach to critical task analysis (SACPATA) was developed during 1985 through applications of human factors engineering and avionics technical skills in the Integrated Facility for Avionics Systems Test (IFAST) facility located at Edwards Air Force Base. SACPATA was used to support the F-16 MSIP and AFTI/F-16 pilot-avionic system task integration analyses.

SACPATA provided a systematic and "forward-looking" process suitable for identifying and modifying pilot-avionic task interface. The SACPATA approach calls for a pilot-avionic task analysis to begin in a conceptual design of the avionic system and continue through development test and operational usage.

Normally, a critical task analysis was accomplished piecemeal using only "after-the-fact" information. Over the years this method worked well because the need for critical task analysis was primarily for training purposes. During those years, most flight test programs were characterized by a "fly-fix-fly" philosophy. The changes and modifications that resulted had little effect on critical task analyses. However, with complex integrated avionic systems featuring computerized master modes and submodes, multi-function/head-up displays with many formats and menu options, it is not "cost-effective" to wait until full-scale development flight tests to initiate an "after-the-fact" critical pilot-avionic task analyses.

In addition, new avionic system designs have become more complicated and problems and difficulties can be encountered during flight tests changing master modes and submodes, setting up display formats, and menu paging to prepare weapons for launch. It is often a real challenge to setup sensor modes and cockpit displays in the time required to accomplish the mission task flow.

It was apparent that something had to be done to resolve critical pilot-avionic task difficulty before flight testing. A systems approach (SACPATA) was drafted that outlined the techniques which could be used during avionic systems development to identify pilot-avionic task choke points and problem areas.

SACPATA approach called for initiating the critical pilot-avionic task analysis in design concept and extending it through developmental stages. SACPATA was characterized by the "identify-analyze-modify" method of pilot-avionic task interfacing. The SACPATA goal during early design was to eliminate or reduce pilot-avionic task problems "before-the-fact" to prevent them from being designed into the system.

In view of the fact that most pilot-avionic system task overloads have been due to avionics system design, the main focus of SACPATA was placed on identifying pilot-avionic system interface problems during development. The pilot-task overloads must be eliminated or reduced to an "acceptable" level prior to flight test and evaluation.

SACPATA should continue through the development and operational phases. The core of the systems approach includes tracking critical pilot-avionic tasks to assure that task problems (hazards) do not "unacceptably" affect mission accomplishment.

2.0 AVIONIC SYSTEM CONCEPT AND DEFINITION PHASES

During system concept and definition phases, pilot-avionic critical task identification and preliminary avionic system design should be accomplished simultaneously. Pilot-avionic task decision trees and mission task flow simulation models can provide a comprehensive investigation of pilot-task choke points and pilot-avionics integration problem areas (hazards) that are associated with the preliminary design. The results of SACPATA should be reported in detail during the concept and definition phase design reviews.

To produce a "problem free" task design that meets the criteria derived in the concept phase, a clear definition of mission critical pilot-avionic task problems must be made. The definition must include "cause and effect" relationship, and the likelihood and severity of the pilot-avionic task problem. If a pilot-avionic task problem persists, alternative system designs must be considered as the first means to eliminate the problem.

To obtain a clear definition of critical pilot-avionic task problems, the mission task flow analysis must be accomplished. The process outlined in the Timebased Analysis of Significant Coordinated Operations (TASCO), "A Cockpit Workload Analysis Technique" document can be used to initially identify the pilot-avionic task overload and potential mission critical task hazards. The timely evaluation of potential mission critical task hazards was a key element in the overall human factors and avionics system integration analysis.

2.1 AVIONIC SYSTEM DEVELOPMENT PHASE

The development phase should include evaluation of pilot-task interfaces through prototype analysis and preliminary design testing.

Since the design at this time is more complete, detailed pilot-avionic system information can be placed into the critical pilot-avionic task analyses.

Precise pilot-avionic system evaluations are required to assess the need and the acceptability of the manual master mode and display format configuration change and the automatic cockpit configuration.

Corrective action must be taken to relieve pilot-task overloads in systems design and the pilot-task workload level "acceptability" must be verified. Close coordination with avionic disciplines should be exercised and specific actions initiated to eliminate or reduce harmful effects of pilot-avionics task overload and/or task hazards on the mission.

Completion of the development phase should lead to a "GO/NO-GO" decision of the final system design before the actual production begins. The analyst's ability to make the correct GO/NO-GO decision will be based upon completion of the preliminary pilot-avionic task analysis and preliminary test results.

2.2 AVIONIC SYSTEM PRODUCTION AND DEPLOYMENT PHASES

An updating of potential task hazards must be continued through the full-scale evaluation and test phase. The comprehensive review of the preliminary pilot-avionic task analysis must be accomplished to verify that appropriate corrective actions have been taken to eliminate or reduce the effects of the task hazard. Modifications and changes in system design to eliminate or reduce potential pilot-avionic task hazards will be subject to review and verification during final design acceptance meetings.

3.0 SACPATA SEQUENCE OF PILOT-AVIONIC TASK HAZARD REMOVAL AND CONTROL

SACPATA will identify and document pilot-avionic task work load levels in the design under development. SACPATA data, resulting from design evolution and evaluation, can play an important role in system acceptance.

SACPATA information must be carried into production and deployment phases to follow-up on potential task hazards that may remain in the system's design or be uncovered through operational usage.

The sequence of pilot-avionic task hazard removal and control should follow a system of prioritized activities. The most preferred activity sequence is as follows:

a. Design to eliminate pilot-avionic task hazards

b. Design to reduce critical task hazards to an acceptable level

c. Control critical task hazards through special task procedures

d. Control critical task hazards through special task training.

Design to eliminate task hazards is an ambitious goal, even though some segments of the mission require strict "design-out" of all pilot-avionic task hazards to avoid an "unacceptable" loss.

Design to reduce task hazards to an "acceptable" level implies a residual problem left in the design of the system. The potential of the task problem may be reduced in magnitude by reduction of its probability and severity. Also, an "acceptable" work load level in one mission segment may be totally "unacceptable" in another mission segment. Task hazard "acceptability" must be weighed in terms of risk benefit factors for each segment of the mission.

Special procedures and training can be used to control a task problem if design methods should fail. In this case, special procedures must be clearly detailed and pilots specially trained. This is not a preferred means of reducing task hazards, but may be the only avenue remaining. Design changes or modifications after definition phase may be complex or costly.

4.0 AVIONIC SYSTEM LIFE-CYCLE

A system's life-cycle is normally separated into distinguishable phases. The most common phases are the concept, definition, development, production, and deployment phases. During these five phases, SACPATA effort should address the following human factors engineering and avionics system integration aspects as shown in Table 1.

Table 1

SYSTEMS APPROACH TO CRITICAL PILOT-AVIONICS TASK ANALYSIS (SACPATA)
HUMAN FACTORS ENGINEERING AND AVIONICS SYSTEM INTEGRATION

Phase	Aspects
Concept	Pilot and Avionic System Interface Conceptual Research
Definition	Critical Pilot-Avionic Task/System Design Verification
Development	Mission Critical Pilot Tasking/Full-Scale Development Verification
Production	Mission Critical Pilot Tasking/Operational Validation
Deployment	Mission Critical Task/Mission Objective Validation

A relative amount of pilot-avionic task analysis (SACPATA) effort that should be expected and planned during each life-cycle phase is shown in Figure 1.

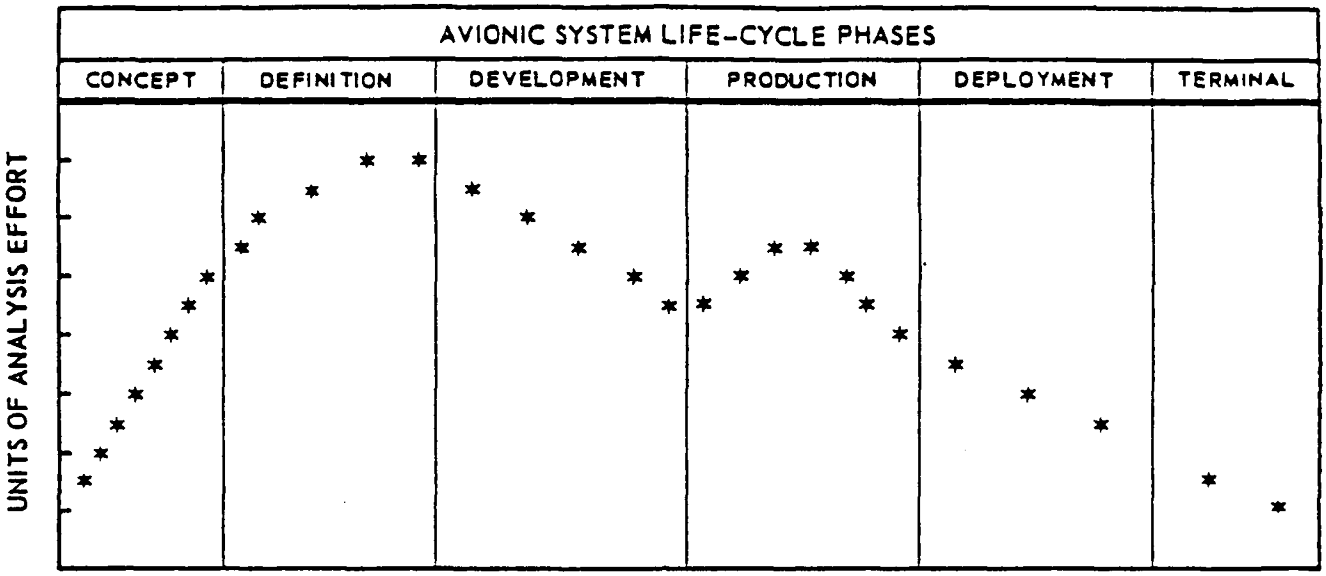


Figure 1 SACPATA Effort During Avionic System Life Cycle

5.0 SACPATA PRIORITY DIAGRAM FOR TASK HAZARD ELIMINATION AND/OR REDUCTION

The priorities that should be used to eliminate or reduce pilot-avionic task hazard by design is shown in Figure 2. If this fails, control task problems through special procedures or special training.

5.1 SACPATA OUTPUTS AND SUMMARY

The outputs from a properly accomplished SACPATA can form a basis for specifying automatic configuration changes over manual control. The data can provide operational and work-around procedures, as well as, risk-benefit factors for a meaningful human engineering and avionics integration analyses.

Expected outputs from critical pilot-avionic task analyses are shown in Figure 3.

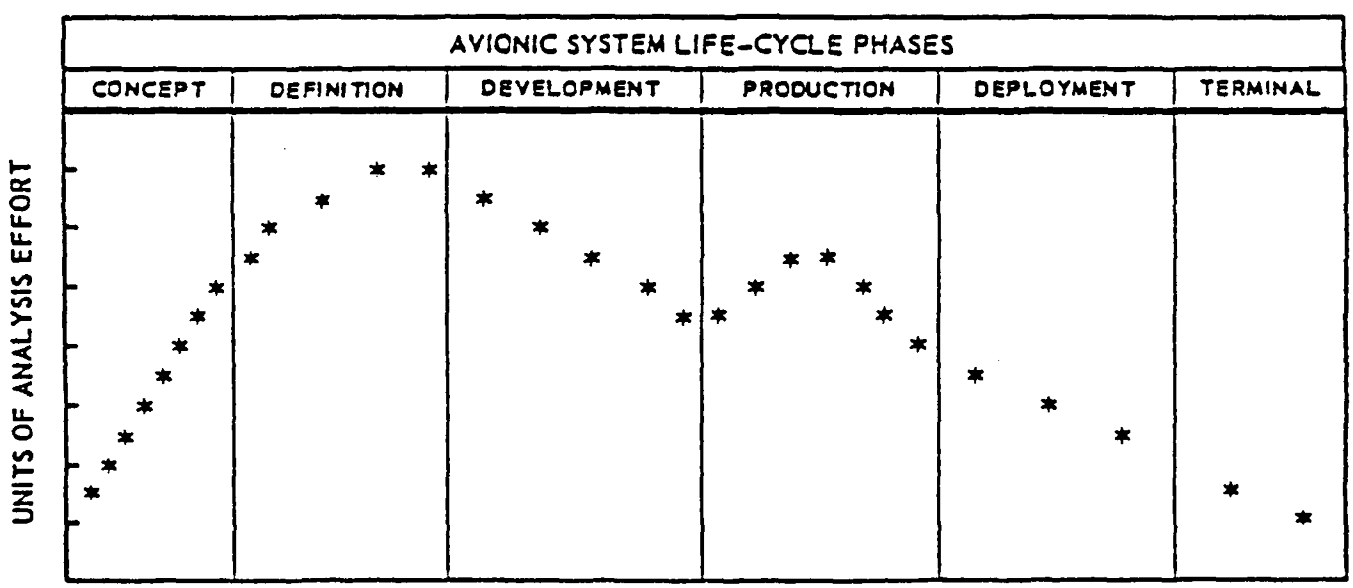
5.2 CRITICAL PILOT-AVIONIC TASK ANALYSIS SUMMARY

The task analysis should not be regarded as a task "failure" analysis. That would be an "after-the-fact" approach. Critical

SYSTEMS APPROACH TO CRITICAL PILOT-AVIONICS TASK ANALYSIS (SACPATA)

HUMAN FACTORS ENGINEERING AND AVIONICS SYSTEM INTEGRATION

Phase	Aspects
Concept	Pilot and Avionic System Interface Conceptual Research
Definition	Critical Pilot-Avionic Task/System Design Verification
Development	Mission Critical Pilot Tasking/Full-Scale Development Verification
Production	Mission Critical Pilot Tasking/Operational Validation
Deployment	Mission Critical Task/Mission Objective Validation



SACPATA Effort During Avionic System Life Cycle

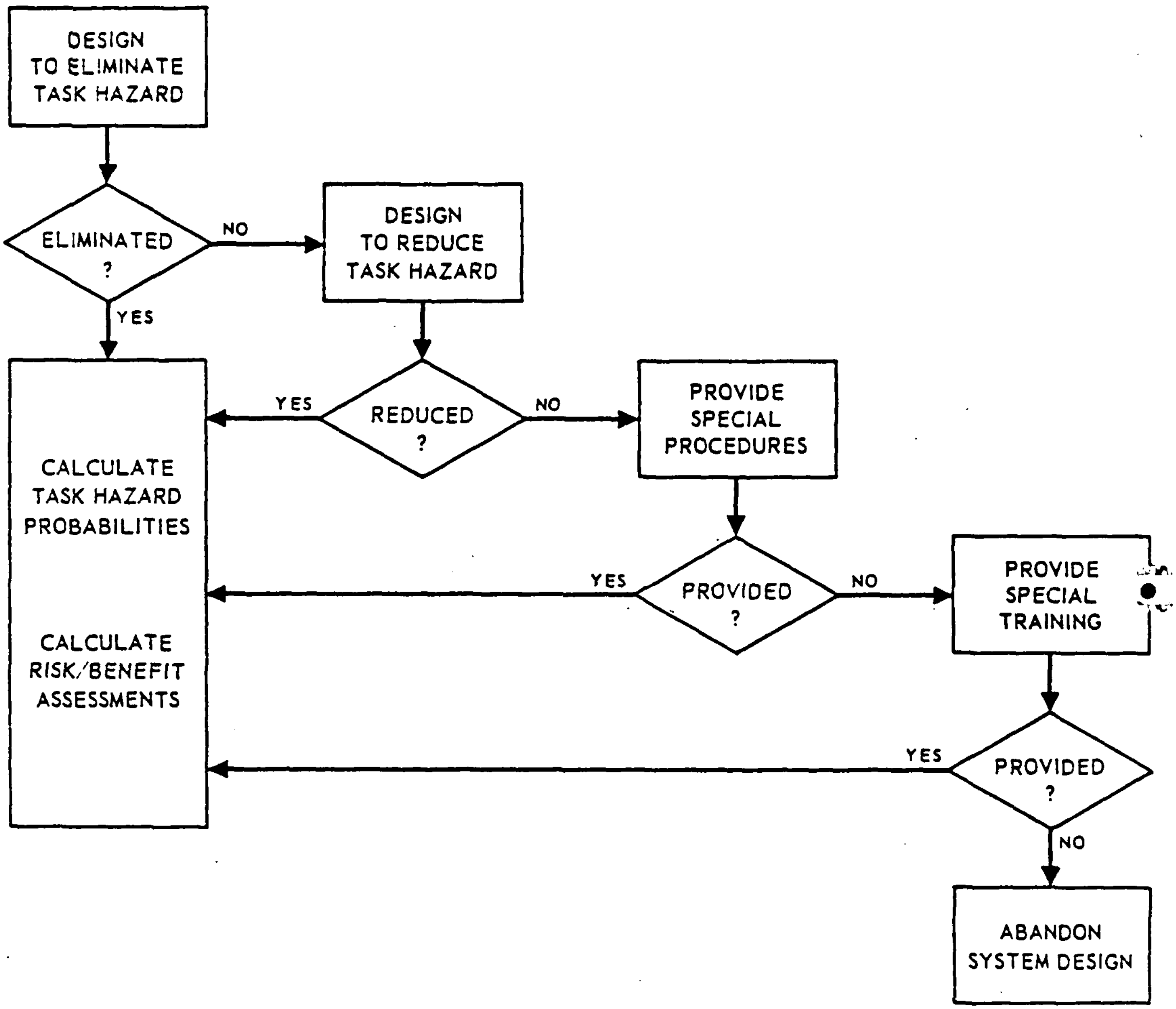
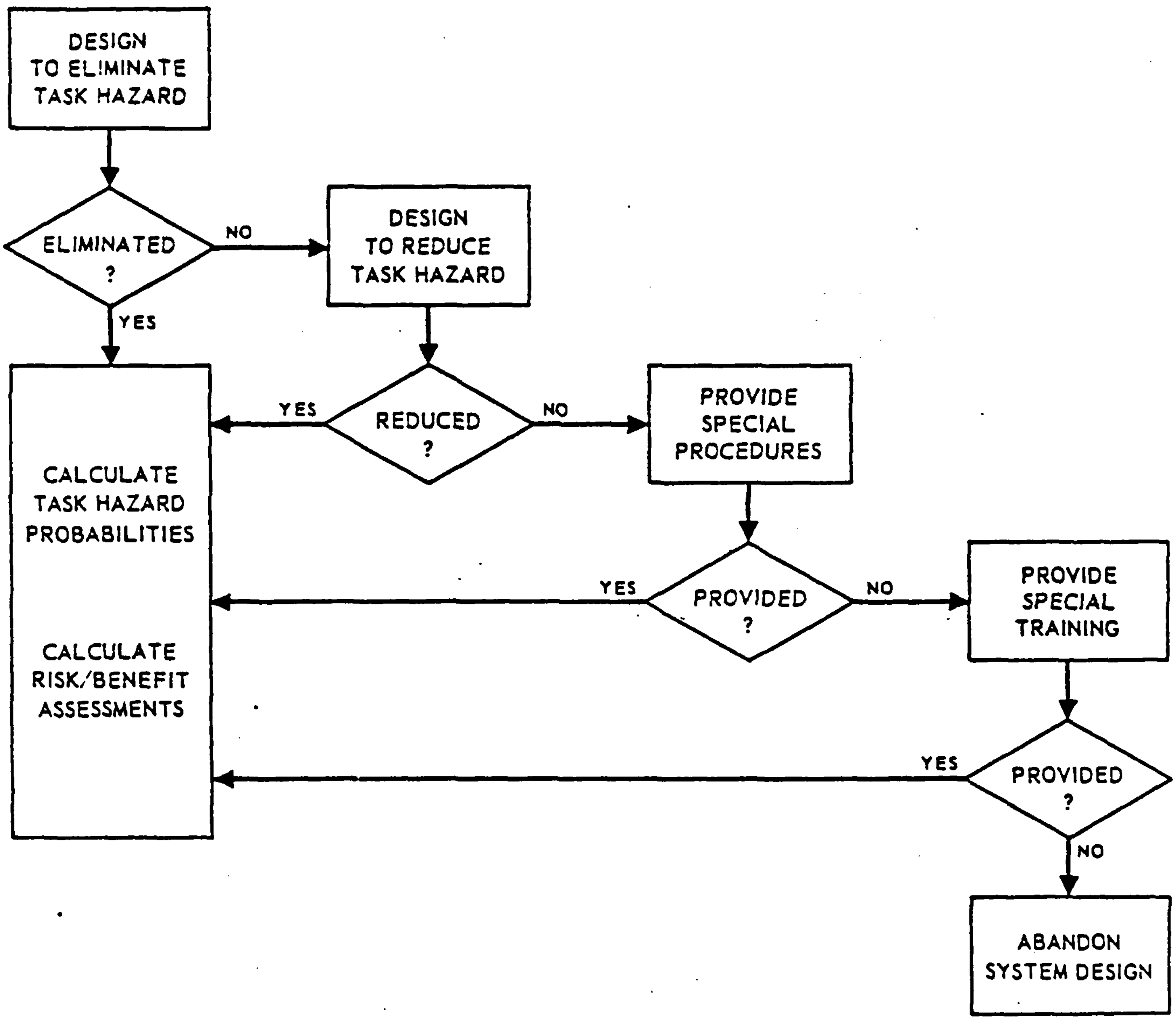


Figure 2 Pilot-Avionic Task Hazard Elimination and/or Reduction



Pilot-Avionic Task Hazard Elimination and/or Reduction

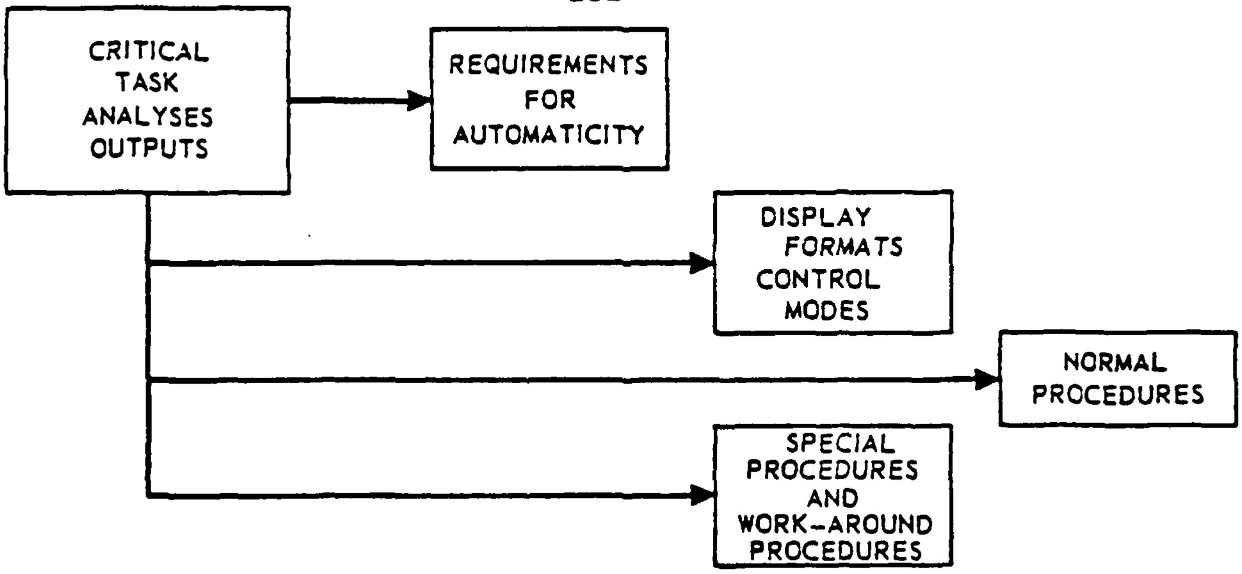


Figure 3 Critical Task Analyses (SACPATA) Outputs

task analysis should occur "before-the-fact" and be regarded as critical task hazard analysis. The analytical methods used are specialized analysis techniques involving probability principles, risk-benefit disciplines, and the logic theory.

Pilot-avionic task procedures must be designed and built-in to an aircraft just as performance, stability, structural integrity, safety, etc. A human factors engineering and avionics integration group should be an important part of an aircraft manufacturer's organization, just as stress, aerodynamics, weight and balance, and armament groups.

Critical task analyses are not unique, but they have not flourished or been formally adopted as an aid in the design of man-machine systems. In fact, critical task analyses concepts have not been clearly defined by Government agencies contracting in the civilian aerospace field. SACPATA represents a first step in establishing critical task analyses definition.

5.3 COMPUTER-AIDED CRITICAL PILOT-AVIONIC TASK ANALYSES

Computer-aided (or computer-based) evaluation and testing in critical task analyses can form vital links between the system preliminary design, objective task flows, and cockpit work load assessment logic structures. A mission task flow evaluation and testing, that was properly conducted, would provide the pilot's inputs to the system model and also serve as an entry point for operator procedures into the task analysis. Figure 4 illustrates a mission critical task flow evaluation and testing process in broad-scale overview.

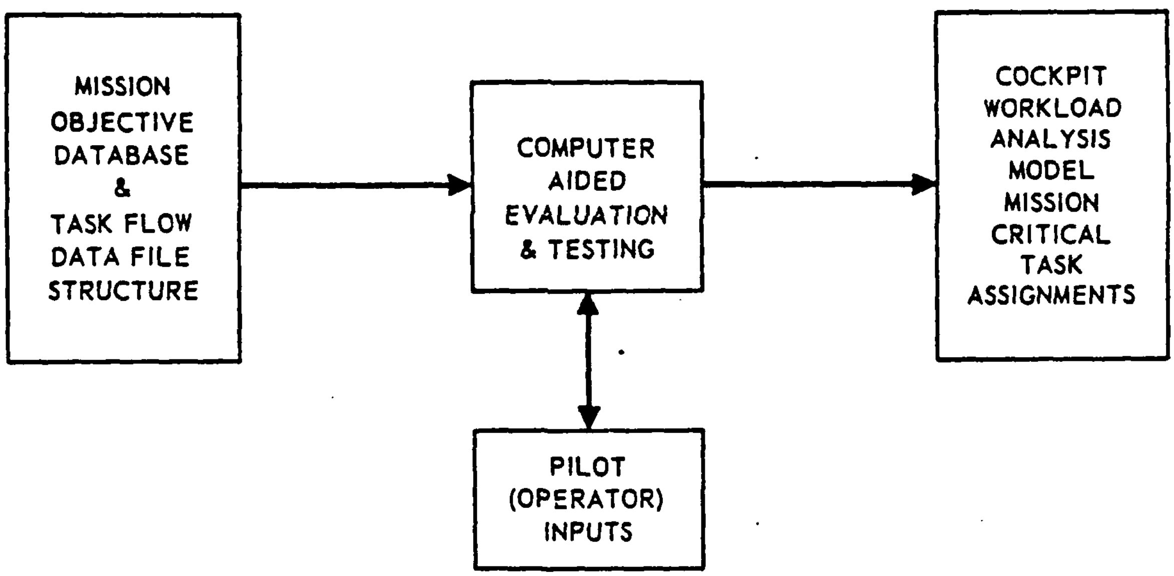
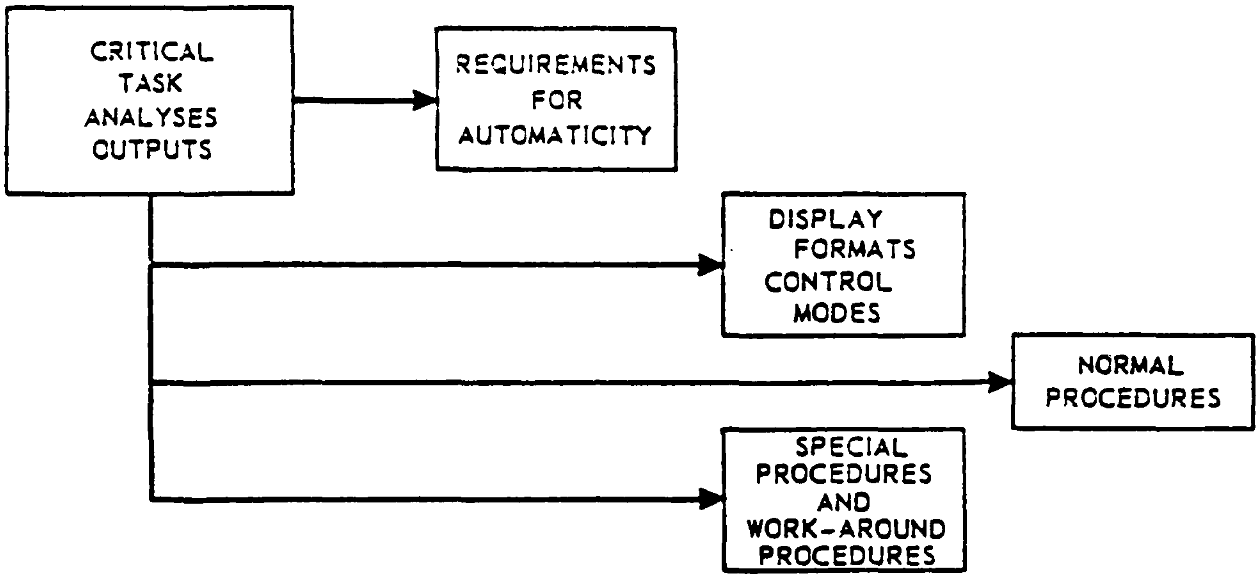
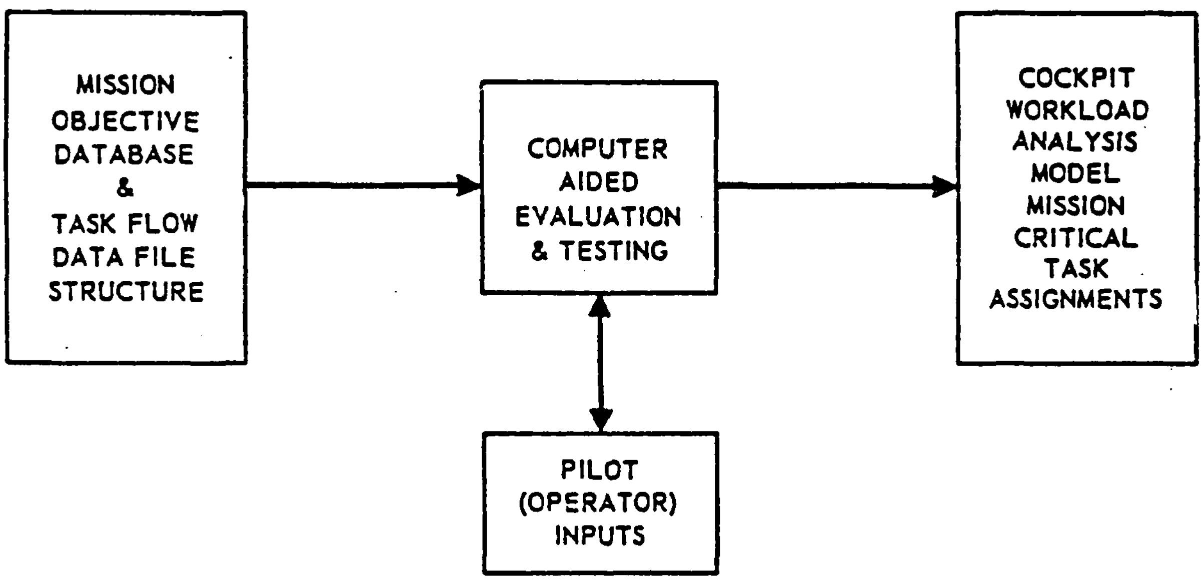


Figure 4 Computer-Aided Mission Critical Task Flow Evaluation/Testing

The mission critical task flow evaluation and testing examines the mission capabilities and objectives with respect to pilot-avionic system capabilities. Aircraft and mission capabilities can be identified early in the concept phase. Preliminary man-machine task allocation can be developed to support the aircraft and mission capabilities. The risk-benefit analysis is used in a preliminary



Critical Task Analyses (SACPATA) Outputs



Computer-Aided Mission Critical Task Flow Evaluation/Testing form to reach conclusions about the importance of the tasking assignments. Three basic questions should be answered at the close of the concept phase.

a. Have the task hazards associated with system design been discovered and evaluated to assure pilot-aircraft-mission objectives have been met?

b. Have risk-benefit analyses been initiated to establish the means of task hazard assessment and control?

c. Are critical pilot-avionic task design requirements established for the concept so that the next phase of system definition can be initiated?

6.0 PILOT-AVIONIC SYSTEM CAPABILITIES AND MISSION OBJECTIVES

When the avionic system is evaluated in the light of mission objectives, strengths and weaknesses surface as pilot-task decision trees are constructed. Pilot opinions and recommendations are interplayed with mission objectives during the evaluation of avionic system capabilities. The SACPATA approach calls for the examination of the total pilot-aircraft-mission system during the early stages of avionic system development. The interplay of pilot, aircraft, and mission is shown in Figure 5.

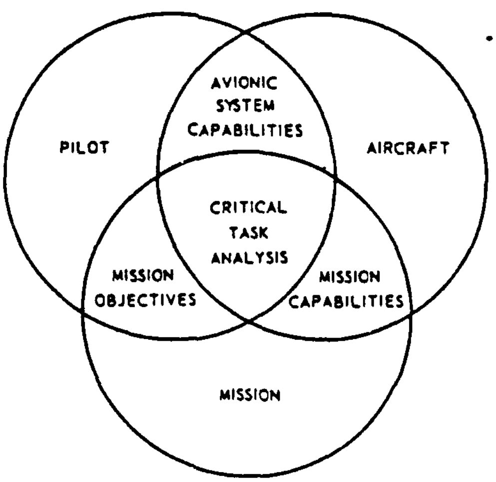
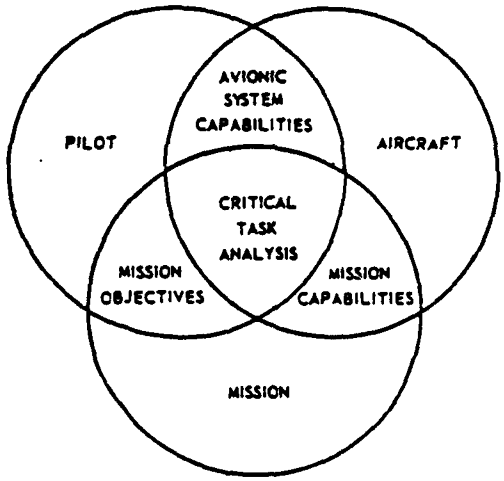


Figure 5 Critical Task Analysis Factors



Critical Task Analysis Factors

In addition to mission critical pilot-avionic task identification, operator evaluation and testing provides many salient products useful in the task analysis which include the following:

a. Strengths and weaknesses of the avionics system surface during the critical pilot-avionic task examination.

b. Alternative techniques and work-around procedures evolve as the pilot-avionic task decision trees are developed.

c. Seldom utilized cockpit procedures and display modes are keynoted.

d. Pass or fail criteria for operator task performance are established.

e. Operator task overloads and mental stress points are highlighted.

7.0 SACPATA COMPUTER-AIDED CRITICAL PILOT-AVIONIC TASK ANALYSES MODEL

The mission scenario is evaluated to produce a "benchmark sortie" having operational characteristics. The benchmark sortie analysis provides cockpit mission critical task assignments through tasking criteria application. The pilot's inputs, through computer-aided evaluation and testing, are examined during cockpit task analysis to determine risk-benefit probabilities for the most critical mission flow pilot-avionic tasks.

Figure 6 shows where and how computer-aided mission critical task flow evaluation and testing fits into the broad-scale pilot-avionic task analysis model.

7.1 CRITICAL PILOT-AVIONIC TASK HAZARD PROBABILITY ANALYSES

Critical task pilot-avionic task analyses can assess the degree of task "hazard" in a system either qualitatively or quantitatively.

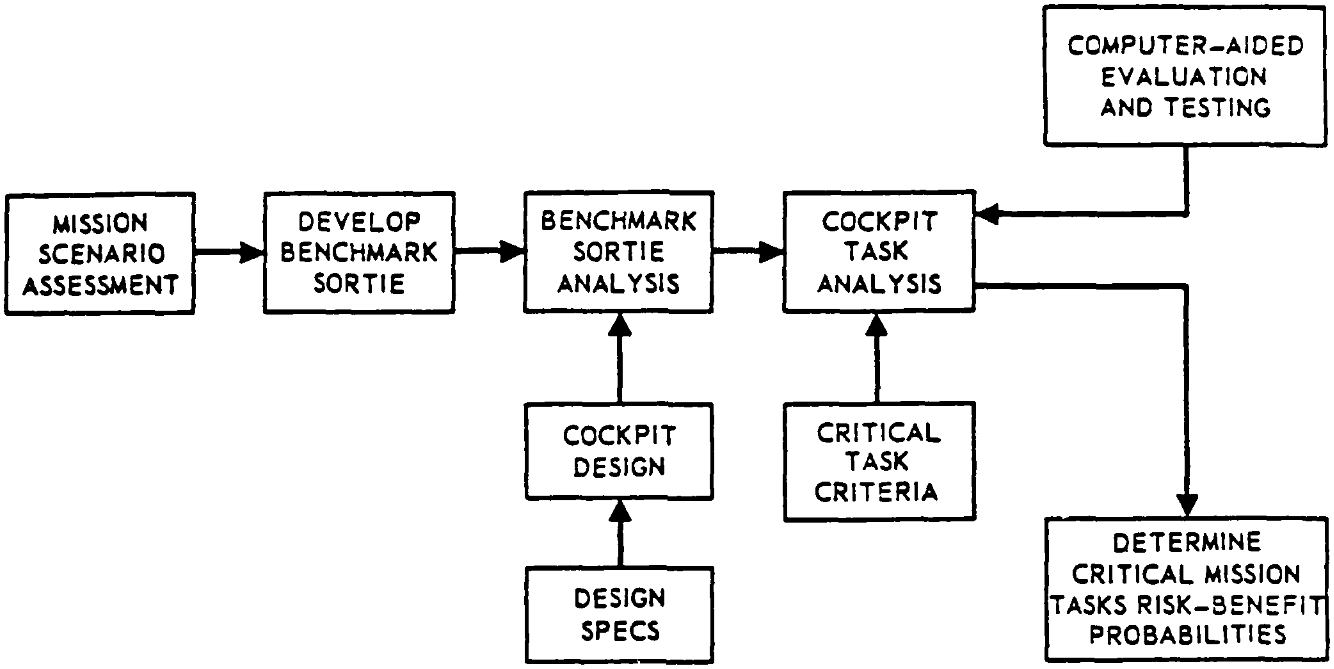
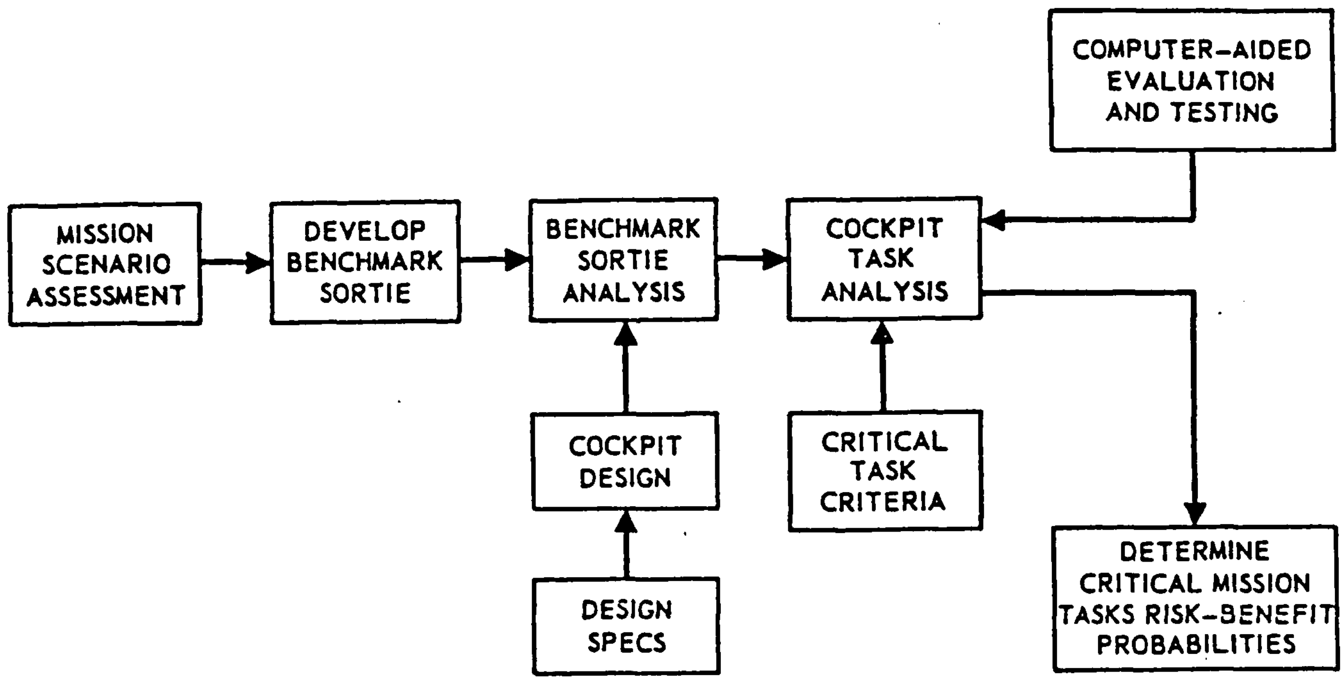


Figure 6 Computer-Aided Critical Pilot-Avionic Task Analyses Model

A qualitative assessment uses the judgemental approach which may lack sophistication and allow biased data to enter the analysis. A quantitative approach uses computerized algorithms requiring the application of rigorous logic for cause-effect relations and probability predictions.

The critical pilot-avionic task "hazard" assessment should be initiated "before-the-fact" and certain statistical methods can be used to account for any uncertainty by probability predictions. There are two basic types of probability statistics. One form, called "a posteriori" probability, is developed by conducting a test or an experiment and observing the outcome. Sample data that are collected during the test are used to derive the probability. The data may be collected from "real-world" events. The following sample aircraft accident data are used for "a posteriori" probability prediction.



Computer-Aided Critical Pilot-Avionic Task Analyses Model

Aircraft Accident Sample Data

5 -	Resulted in death
10 -	Resulted in severe injury
30 -	Resulted in moderate injury
75 -	Resulted in minor injury
250 -	Resulted in no injuries

370	Total number of accidents

Probability of Death (d), given that one is in an aircraft accident, is:

$$P(d) = 5/370 = 0.0135$$

An "a priori" probability can be specified by the evident nature of the events from which they emerge. The "a priori" probability of a head (h) coming up from a toss of a fair coin can be calculated as 0.5.

$$P(h) + P(t) = 0.5 + 0.5 = 1.0$$
$$P(h) = 0.5$$

The critical pilot-avionic task analysis uses "a priori" probabilities for making decisions concerning the acceptability of pilot-task "hazard."

7.2 MISSION CRITICAL PILOT-AVIONIC TASK HAZARD

A "hazard" may be simply defined as "a potential for doing harm." The "harm" in a pilot-avionic task is not being in the correct sensor mode for taking a navigation fix which results in a missed way point, or not being in the right weapon mode for launching a missile and missing a target. The word "potential" is important to the definition of the critical task hazard as a task analyst must control them to an "acceptable" level by using one of several means.

Task hazards possess two inherent properties; likelihood and severity. If either of these two properties has no value, then the task hazard has no potential for harm and is intrinsically safe. Whenever a critical task hazard is found through the task hazard analysis process that is not intrinsically safe, it must be evaluated for acceptability in an operational mission environment. If the task hazard is "unacceptable," it must be dealt with by system design changes, modifications, or controlled through countermeasures such as work-around procedures.

A system of critical task hazard severity can be setup similar in form to the four main categories used in Table 2. These categories are used in the aerospace industry to deal with safety hazards.

Table 2

SAFETY HAZARD CLASSIFICATION

Category	Name	Characteristics
I	Catastrophic	Death - loss of system
II	Critical	Severe injury - major damage
III	Marginal	Minor injury - minor damage
IV	Negligible	No injury - no damage

As mentioned previously, critical task hazard probability refers to the probability occurrence of an event (or the non-occurrence of an event), that results in harm. In the case of mission task flow, the "harm" may be due to being out-of-mode or not in the appropriate display format to complete the mission task.

Early in the system design stage, there may be insufficient information to perform precise critical pilot-avionic task analysis.

However, as the system matures and information is available, probability evaluation can be accomplished with more precision.

7.3 CRITICAL TASK HAZARD UNDESIRABILITY AND SEVERITY CLASSIFICATION

A critical task hazard measure of undesirability (or disutility) can be related to its level of severity and probability. Table 3 illustrates relative degree of disutility of task hazards as a function of probability and severity.

Table 3

CRITICAL TASK HAZARD (DISUTILITY) UNDESIRABILITY

Probability	Severity	
	LOW	HIGH
LOW	Low-Low (low)	Low-High (moderate)
HIGH	High-Low (moderate)	High-High (high)

The least "undesirable" task hazard is in the upper left corner and has a low value (Low-Low) of disutility. The lower right corner has the highest relative value of disutility (High-High) and is the most "undesirable" task hazard.

The critical pilot-avionic task "hazards" are similar to safety hazards. A severity classification can be applied to pilot-avionic task hazards, but the categories must be expanded for a broader-range "harm potential" to the mission. Table 4 shows a "broader-scale" severity classification used for the critical pilot-avionic task hazard severity classification.

SAFETY HAZARD CLASSIFICATION

Category	Name	Characteristics
I	Catastrophic	Death - loss of system
II	Critical	Severe injury - major damage
III	Marginal	Minor injury - minor damage
IV	Negligible	No injury - no damage

CRITICAL TASK HAZARD (DISUTILITY) UNDESIRABILITY

Probability	Severity	
	LOW	HIGH
LOW	Low-Low (low)	Low-High (moderate)
HIGH	High-Low (moderate)	High-High (high)

Table 4

CRITICAL PILOT-AVIONIC TASK HAZARD SEVERITY CLASSIFICATION

Severity Index	Effect on Mission
6	No effect on mission or aircraft
5	Minor effect-(missed navigation waypoint)
4	Major effect-(missed target initial point)
3	Major effect-(missed target of opportunity)
2	Major effect-(missed mission objective target)
1	Major effect-(aircraft damage or loss)

7.4 CRITICAL PILOT-AVIONIC TASK HAZARD QUALITATIVE PROBABILITY RANKING

The qualitative critical task hazard probability ranking provides rating scales for occurrence of task hazards and is used in risk-benefit definition to determine those areas where task modification must be accomplished. In some areas, automatic change of master mode and submode, and display configuration priority have been used to reduce or control "unacceptable" task hazards. In cases of automatic mode changes (pre-assigned display formats and sensor modes) an evaluation must be made to determine their actual impact on the pilot's mission task flow. Increases in automaticity can reduce the pilot's flexibility, do harm to his mission capability, or result in an unacceptable system configuration in another segment of the mission.

Table 5 provides a critical pilot-avionic task hazard qualitative probability ranking that can be used in the SACPATA.

Table 5

CRITICAL PILOT-AVIONIC TASK HAZARD QUALITATIVE PROBABILITY RANKING

Occurrence	Probability	Test Experience	Operational Usage
Frequently probable	6	Likely to occur frequently	Likely to be experienced continuously
Reasonably probable	5	Will occur several times	Will occur many times
Occasionally probable	4	Likely to occur several times	Will occur several times
Remotely probable	3	So unlikely, its not expected	Unlikely to occur, but possible
Extremely improbable	2	Occurrence times next to zero	Occurrence only after other complications
Impossible to occur	1	Physically impossible	Physically impossible

7.5 CRITICAL PILOT-AVIONIC TASK HAZARD INDEX

A critical pilot-avionic task hazard index can be developed from the likelihood (probability) and the undesirable (severity) effects. This task hazard index can then be used to differentiate the effect of the task hazard on the mission.

Figure 7 uses the probability (likelihood) and severity (undesirable effects) indices to develop "critical" pilot-avionic task hazard indices for differentiating harmful effects on the mission.

The "critical" pilot-avionic task hazard index is the product of probability and severity points. Having formed the basic matrix of critical task hazard indices, the analyst can designate the level

CRITICAL PILOT-AVIONIC TASK HAZARD SEVERITY CLASSIFICATION

Severity Index	Effect on Mission
6	No effect on mission or aircraft
5	Minor effect-(missed navigation waypoint)
4	Major effect-(missed target initial point)
3	Major effect-(missed target of opportunity)
2	Major effect-(missed mission objective target)
1	Major effect-(aircraft damage or loss)

CRITICAL PILOT-AVIONIC TASK HAZARD QUALITATIVE PROBABILITY RANKING

CRITICAL PILOT-AVIONIC TASK HAZARD QUALITATIVE PROBABILITY RANKING

Occurrence	Probability	Test Experience	Operational Usage
Frequently probable	6	Likely to occur frequently	Likely to be experienced continuously
Reasonably probable	5	Will occur several times	Will occur many times
Occasionally probable	4	Likely to occur several times	Will occur several times
Remotely probable	3	So unlikely, its not expected	Unlikely to occur, but possible
Extremely improbable	2	Occurrence times next to zero	Occurrence only after other complications
Impossible to occur	1	Physically impossible	Physically impossible

CRITICAL PILOT-AVIONIC TASK HAZARD INDEX

PROBABILITY INDICES	SEVERITY INDICES					
	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	8	10	12
3	3	6	9	12	15	18
4	4	8	12	16	20	24
5	5	10	15	20	25	30
6	6	12	18	24	30	36

Figure 7 Critical Pilot-Avionic Task Hazard Indices

of task hazard index that is "unacceptable" or "critical" for that segment of the mission (e.g.,10), as shown by the dotted line in Figure 7.

8.0 CRITICAL PILOT-AVIONIC TASK DECISION TREE ANALYSIS METHOD

The decision tree is fundamentally a Boolean logic model that depicts the relationship between events in a system that lead to a final outcome event. The final outcome event of the pilot-avionic task tree intersects a mission task flow at the head (top-level) event. Subelement events are below the head event and form logical occurrences required to achieve final outcome of the head event.

The pilot-avionic task decision tree and mission task flow intersection, as used in SACPATA, is depicted in Figure 8.

While the method carries in the title "Decision Tree," connoting a Top-Level Head Event as a decision, the tree may be constructed with the top-level event such as the one depicted in Figure 8. In that case, the top-level event was "Verify AA/LCOS Mode/Format." The tree would then become the pilot-avionic system interactions required to produce or satisfy the top-level (head) event.

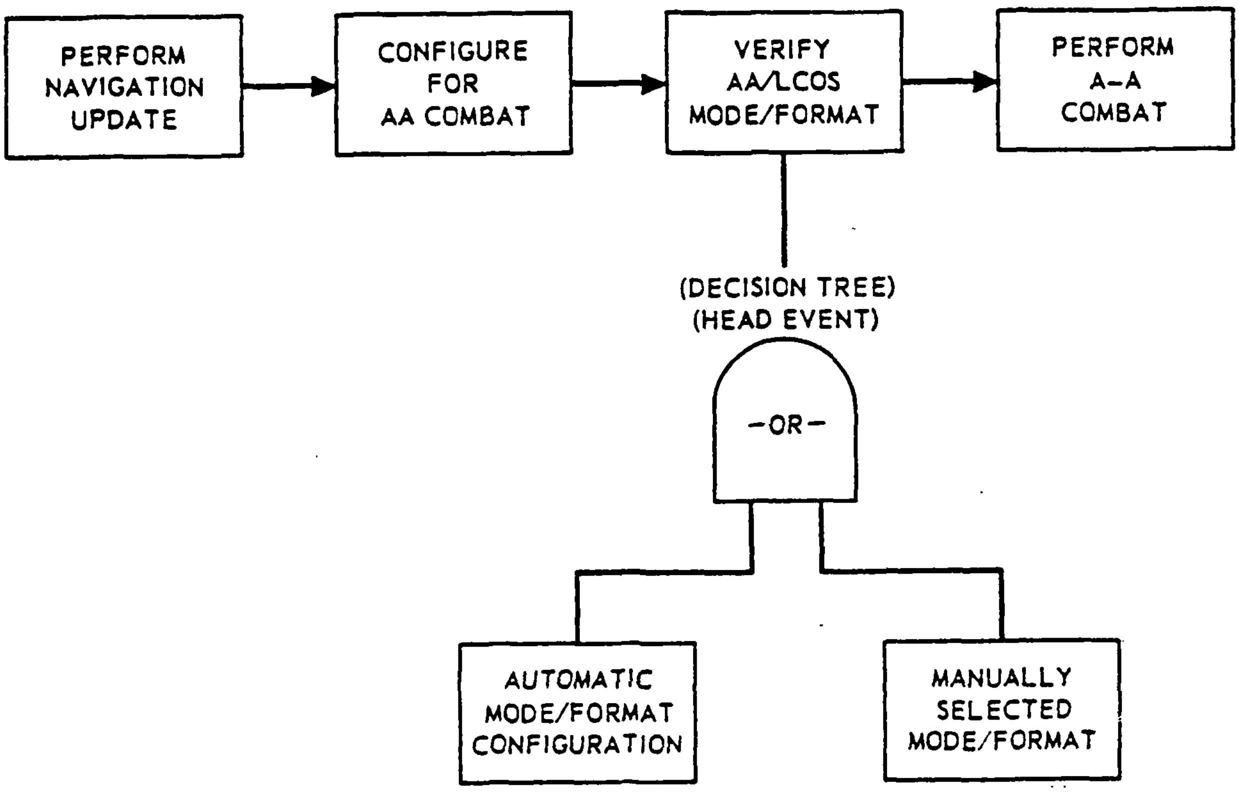
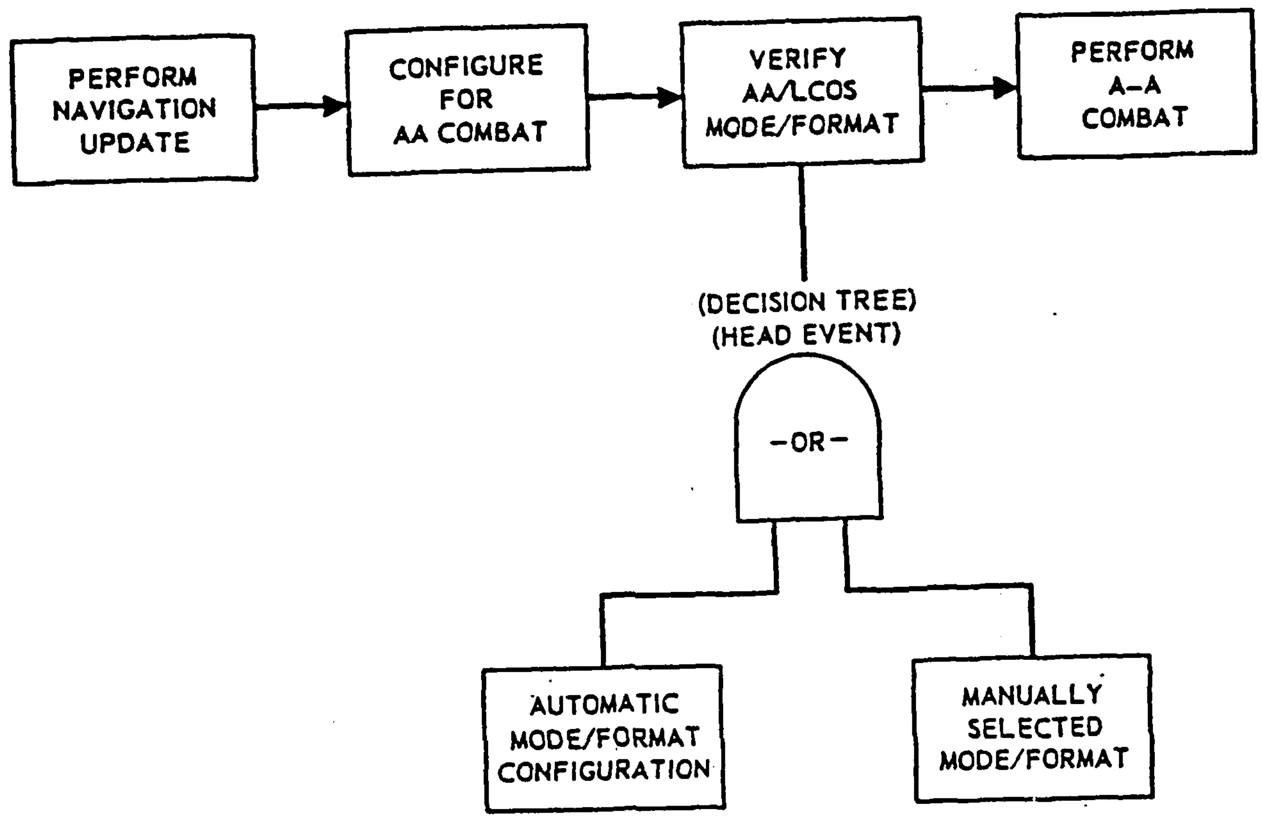


Figure 8 Mission Task Flow and Critical Pilot-Avionic Task Intersection

8.1 PILOT-AVIONIC TASK DECISION TREE (SACPATA) APPROACH

The task decision tree approach is a systematic, descriptive form of analysis that may be applied to task analyses. It can be useful in the early design phase of a new system and particularly important for analyzing operational systems. The decision tree method allows the analyst to evaluate alternatives and judge acceptable trade-offs among them. The decision tree method has the power



Mission Task Flow and Critical
Pilot-Avionic Task Intersection

of deduction where combination subelement events should be considered in the causal chain. The interactions between events and subsystems form vital parts for understanding how the system functions, or how it should function.

As mentioned previously, the analysis of a system through a decision tree can result in either a qualitative or a quantitative output, or both whichever is most advantageous for the analyst. The qualitative output results in sets of events that effect the top-level event and ranking of each event for sensitivity of the top-level event to its occurrence or importance.

The analyst may obtain quantitative probabilities or rates of occurrence of the basic events to obtain a more precise ranking of the importance of each event, as well as other probalistic measures.

The pilot-avionic task decision tree is prepared by a computerized function that diagrams contributing factors linked through logic gates to a head event. The task decision tree clearly, precisely, and concisely defines the top-level mission flow head event for which nonoccurrence would be critical to the mission task flow.

Each subelement of the event that is capable of producing an occurrence or adding to the capability of producing the occurrence of the event is examined. Then a determination is made as to how its failure to occur could contribute to the failure of the head event to occur. When more than one subelement on a task tree could contribute to the same effect on the event, a determination is made as to whether input subelements must act in combination (-AND- relationship) to produce the effect, or whether they may act singly (-OR- relationship) to produce the same effect. The relationships for the pilot-avionic task decision tree are simplified, as much as possible, to ease construction.

Mathematical expressions for representing pilot-avionic task decision trees are developed using simple Boolean algebra. Possibility of each subelement via a logic model producing nonoccurrence of a head event is determined by probability density and distribution functions.

As data becomes available from simulations and flight tests designed to examine the mission flow tasking, the basic function probabilities should be modified or appropriately adjusted to reduce risk for nonoccurrence of the head event.

8.2 PILOT-AVIONIC TASK DECISION TREE ASSUMPTIONS AND STIPULATIONS

Additional fallouts which may result from using the task decision tree are:

a. Identification of critical mode changes requiring automaticity.

b. Discovery of hidden or latent single-point failure probabilities.

c. Determination of the most critical and most probable sequence of events that could lead to an out-of-mode condition.

The pilot-avionic task decision tree is basically a logic diagram that shows various subelements that culminate in the predetermined head event occurrence or produce nonoccurrence. Certain preliminary assumptions and stipulations should be made concerning the characteristics, conditions, and actions involving the head event and subelements. These assumptions and stipulations are stated as follows:

a. Basic subelement events producing the same effect on the head event must be independent of each other.

b. Head events and subelement events have only two conditional modes - totally operative or "in-mode", or totally non-operative or "out-of-mode." No partial operation or partially "in-mode" is included in the analysis.

c. Risk-benefit analyses must use "worse-case" severity indices for nonoccurrence of each top-level or subelement event.

As mentioned before, the analyst may obtain quantitative probabilities or rates of nonoccurrence of basic events. He may also obtain a more precise ranking of the importance of each event, as well as other probabilistic measures.

8.3 PILOT-AVIONIC TASK DECISION TREE SYNTHESIS

Task decision tree synthesizing proceeds fundamentally by the analyst's repeated questions asking what are the "real" decisions that have to be made for each of the basic events. It is obvious that the analyst must not only have a thorough understanding of the avionic system, but be well versed in related areas of knowledge that generate human factors engineering concerns.

Human capabilities and performance must be carefully analyzed with system interfaces. Intimate knowledge of the avionic system requires long sessions with system designers and operators to create a full understanding of design philosophy and methods of system operation. If subsystems are developed and provided by another organization or company, information may be difficult to obtain. Much of the value of the decision tree development has to come from intimate understandings as the analyst probes various designers or design teams. There may have been little, or no prior communication between design teams or designers of individual subsystems. The task analyst must bridge this gap.

In an attempt to simplify the design solutions, a system designer usually will separate the design into subdesign problems. Such an approach can leave interactive gaps for understanding full system functions. The pilot-avionic task decision tree structure with enforced analytical processes and expanded viewpoint of the analyst should uncover many oversights in system design.

While computer codes are used extensively in analyzing the pilot-avionic task decision tree after it has been developed, the intercourse between the designers and analyst is the point at which the decision tree is synthesized. Unfortunately, there is no computer-synthesized task decision tree at the present time.

In addition to those advantages previously mentioned, a properly synthesized task decision tree can provide the following advantages:

a. Direct the analyst deductively to mission critical pilot-avionic task hazard events.

b. Provide a graphical depiction of operator and avionic system interface functions most critical to mission accomplishment.

c. Provide options for both qualitative and quantitative critical pilot-avionic task analyses.

d. Provide the analyst with an insight to the avionics system's function and behavior.

The pilot-avionic task decision tree synthesis process follows prescribed and finely detailed elemental considerations. It forces the analyst to further understand the system beyond the level of the designer of any subsystem or component. Once the logic of the tree has been assembled, intermediate steps for nonoccurrence of mission-critical pilot-avionic task events become readily apparent to the task analyst.

8.4 PILOT-AVIONIC TASK DECISION TREE CONSTRUCTION

The construction of the decision tree usually commences with the top-level (head) event and proceeds downward through successive levels of intermediate effect steps. The analyst must determine at each sublevel of events what the next lower set of events are and if they are both sufficient and necessary to reach events at the base level.

A top-level (head) event is extracted from the mission task flow analysis. In addition to determining the top-level event as origin in the development of the pilot-avionic task decision tree, the analyst must determine the state of the system at the time it is analyzed for occurrence of the top event.

For example, if the analyst should choose "Verify A-A Combat Configuration" as a top-level (head) event, the decision tree would be different depending upon whether the "A-A Configuration" would mean gun-firing or missile firing.

8.5 PILOT-AVIONIC TASK DECISION TREE SYMBOLOGY

As the analyst goes down the pilot-avionic task decision tree from event to event determining "what the fundamental decisions are," logic gate symbols are used to show the logical relations among events. The logic gate symbols have been simplified or minimized for aiding construction of the pilot-avionic task decision tree.

Logic gate symbols depict the logical relations among events of the tree. There are two basic logical relations with several subsets for each. The logical -OR- gate is a logical relation that requires occurrence of one of the input events to effect occurrence of the output event. In most systems, if one subelement event occurs, others are excluded from occurring. This form of relationship, through a logic gate is referred to as an exclusive -OR-. The task decision tree uses the exclusive -OR- gate form.

The other principal logic gate is the -AND- gate which requires that all of the input events must occur to effect the output event occurrence. A restriction that may be placed on an -AND- gate is a sequence requirement that requires the input events to occur in an order to effect the output occurrence. This is called a priority modification of the -AND- gate.

Top level and subelemental events are shown on the tree as rectangles. If the analysis is to be quantitative, a basic event will be assigned a rating or a probability, and the top-level event will be determined by Boolean reduction of the tree.

8.6 PILOT-AVIONIC TASK DECISION TREE ANALYTICAL PROCESS

Once logically constructed, the task decision tree depicts a Boolean model of the system. Having constructed the system model, it must be exercised or analyzed if useful information is to be gained. The visual examination of the pictorial diagramatic description of the task decision tree may be of some use for identifying task choke points. However, substantial information is lost by just a cursory examination of the critical task processes. Information that is available, only from full analyses, points the way to a detailed statement of the system's present state and recommended modifications for improvement.

8.7 BOOLEAN EQUIVALENT TREE

Having developed the Boolean relationships for two fundamental logic gates (-AND-/-OR-), a Boolean equivalent expression can be developed for each of the task decision trees. If the basic events are independent and all of the basic events must occur to allow the top-level event to occur, a multiplication rule applies as shown in Figure 9.

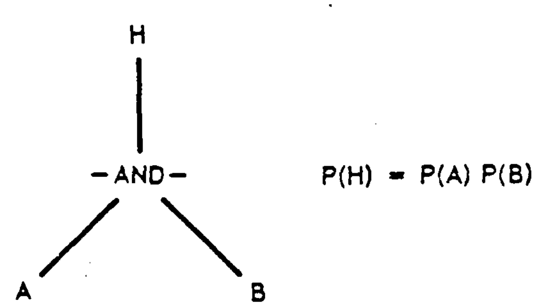
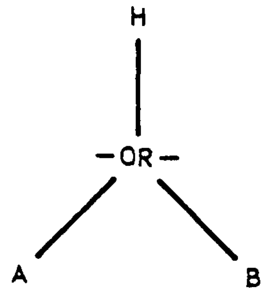


Figure 9 -and- Gate

If the basic events are mutually exclusive, then the rule of addition applies as shown in Figure 10.

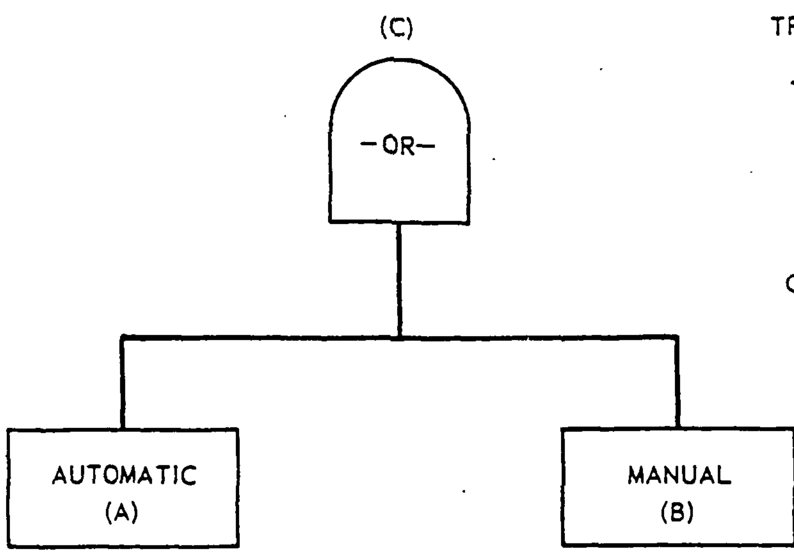


$$P(H) = P(A) + P(B)$$

Figure 10 -or- Gate

8.8 PILOT-AVIONIC TASK DECISION TREE ABSOLUTE LOGIC TRUTH TABLE

A truth table can be developed for the -OR- and -AND- gates which is used to define the input and output words. Figures 11 and 12 show the truth tables for the logical gates used in the critical task decision tree.



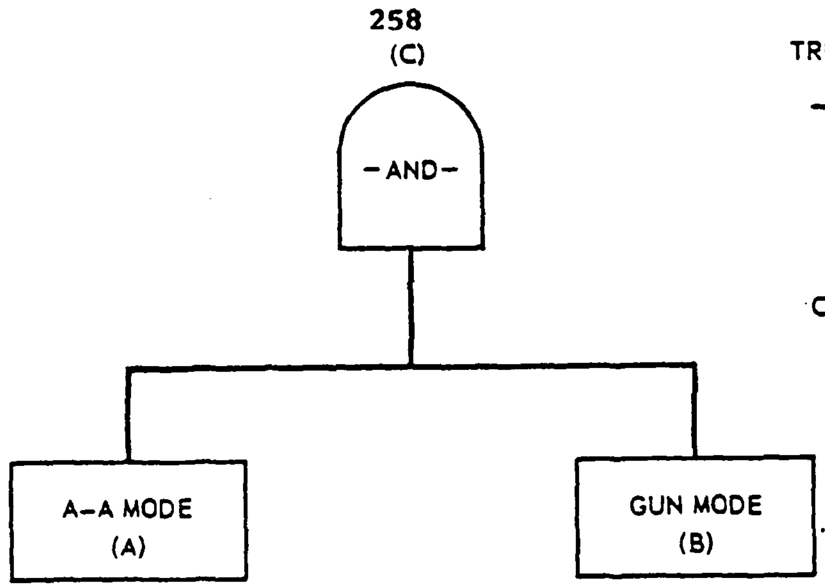
TRUTH TABLE

A	B	C
0	0	0
0	1	1
1	0	1
1	1	1

$$C = A + B$$

Figure 11 Logical -or- Truth Table

The exclusive -OR- gate logic involves the input words (AB), "00,10,01" and the output words (C), "0, 1, 1."



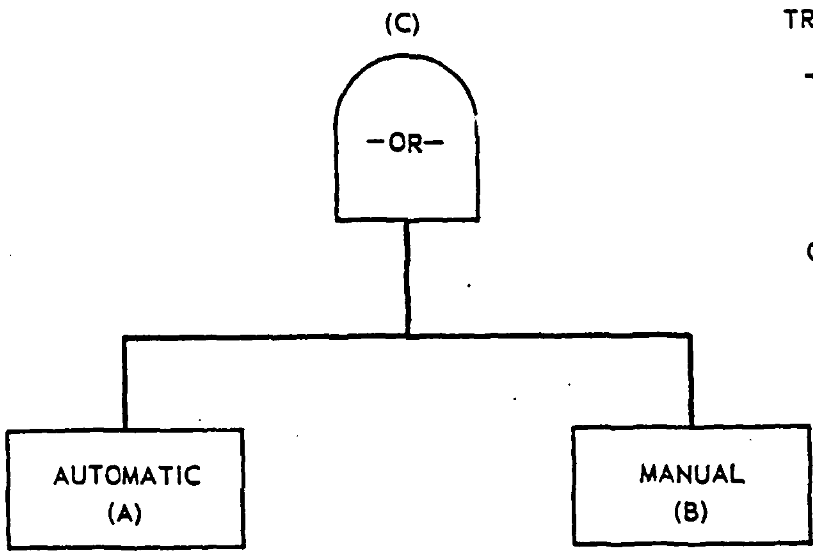
TRUTH TABLE

A	B	C
0	0	0
0	1	0
1	0	0
1	1	1

C = A + B

Figure 12 Logical -and- Gate Truth Table

The -AND- gate logic involves input words (AB), "00, 01, 10, 11" and output words (C), "0, 0, 0, 1." Figure 13 shows the truth table for the logical gates used in the critical task decision tree.



TRUTH TABLE

A	B	C
0	0	0
0	1	1
1	0	1

C = A + B

Figure 13 Logical -or- Truth Table

The exclusive -OR- gate logic involves the input words, (AB), "00, 10, 01" and the output words (C), "0, 1, 1."

9.0 PILOT-AVIONIC TASK MISSION RISK AND BENEFIT ASSESSMENT

There should be an optimal balance between risks and benefits in every mission. However, before an optimal balance can be arrived at, there must be a risk and benefit evaluation to the degree of accuracy required by the parameters of the decision process.

The principle of risk as a potential loss is conceptually simple; however, the assessment of risk is complex due to the number of comprizing elements. A risk assessment model should have the following minimum elements as shown in Figure 14.

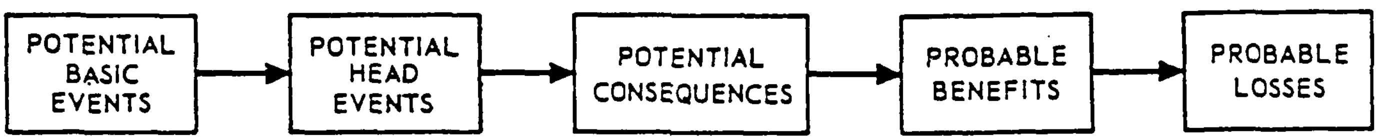


Figure 14 Risk and Benefit Assessment Model

Potential basic events are low-order, causal events that may prevent the occurrence of the head event. Potential head events are a set of top-level mission task flow events. The implication for risk assessment is that one must consider a number of top-level events that may be involved. A potential head event may be thought of as the critical event for the accomplishment of the mission task flow. Basic events include all of the man-machine interface actions required to accomplish the head event.

The potential consequences are the immediate results of the nonoccurrence of the head event. These consequences are dependent upon mission circumstances and are highly variable. However, potential consequences can describe the achievement losses.

The probable benefits place a value on mission accomplishment. Unfortunately, there is always a probable loss to the mission achievement, such as missing a navigation way point. The assessment raises a number of questions such as, how to determine the potential loss to the mission by missing a navigation way point, or by missing a target of opportunity? If the severity of the potential loss is great, such as missing a mission objective target or by loss of an aircraft, then the assessment is less difficult.

In the final analysis, judgement as to "acceptability" of risk is always highly subjective. However, the criteria required to set subjective decision points must be objective in nature and present information to help balance the benefits against the risks.

Table 6 illustrates the critical task probability-risk-severity relationship that should be considered in the risk and benefit assessment.

Table 6

PROBABILITY-RISK-SEVERITY RELATIONSHIP

Probability	Risk	Severity
1. Impossible	6. None	6. Nill
2. Improbable	5. Very small	5. Nuisance
3. Small	4. Small	4. Marginal
4. Significant	3. Significant	3. Significant
5. High	2. High	2. Extreme
6. Extreme	1. Extreme	1. Catastrophic

Figure 15 shows the relationship of probability, risk, and severity in diagramatic form. By cross referencing Table 6 and Figure 15, it should be noted that when the risk is extreme (1) and the severity is catastrophic (1), the probability is impossible (1).

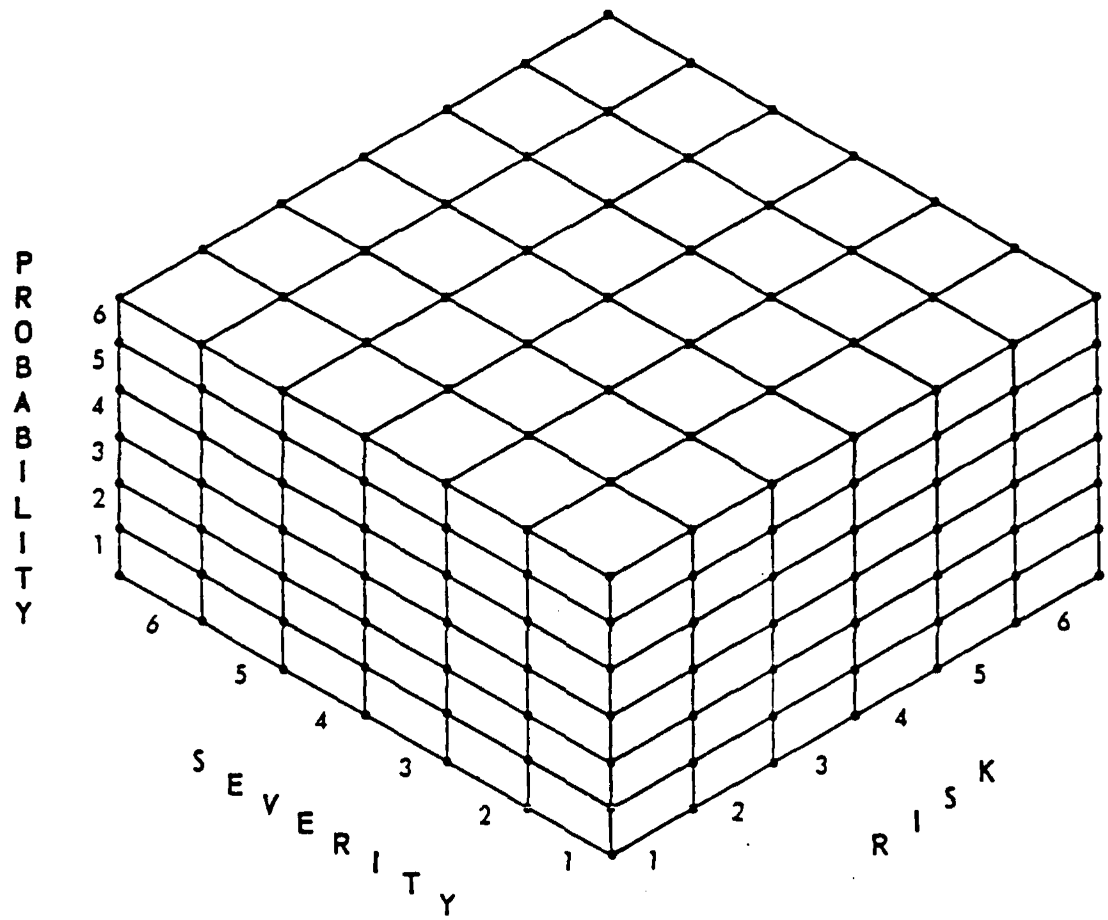
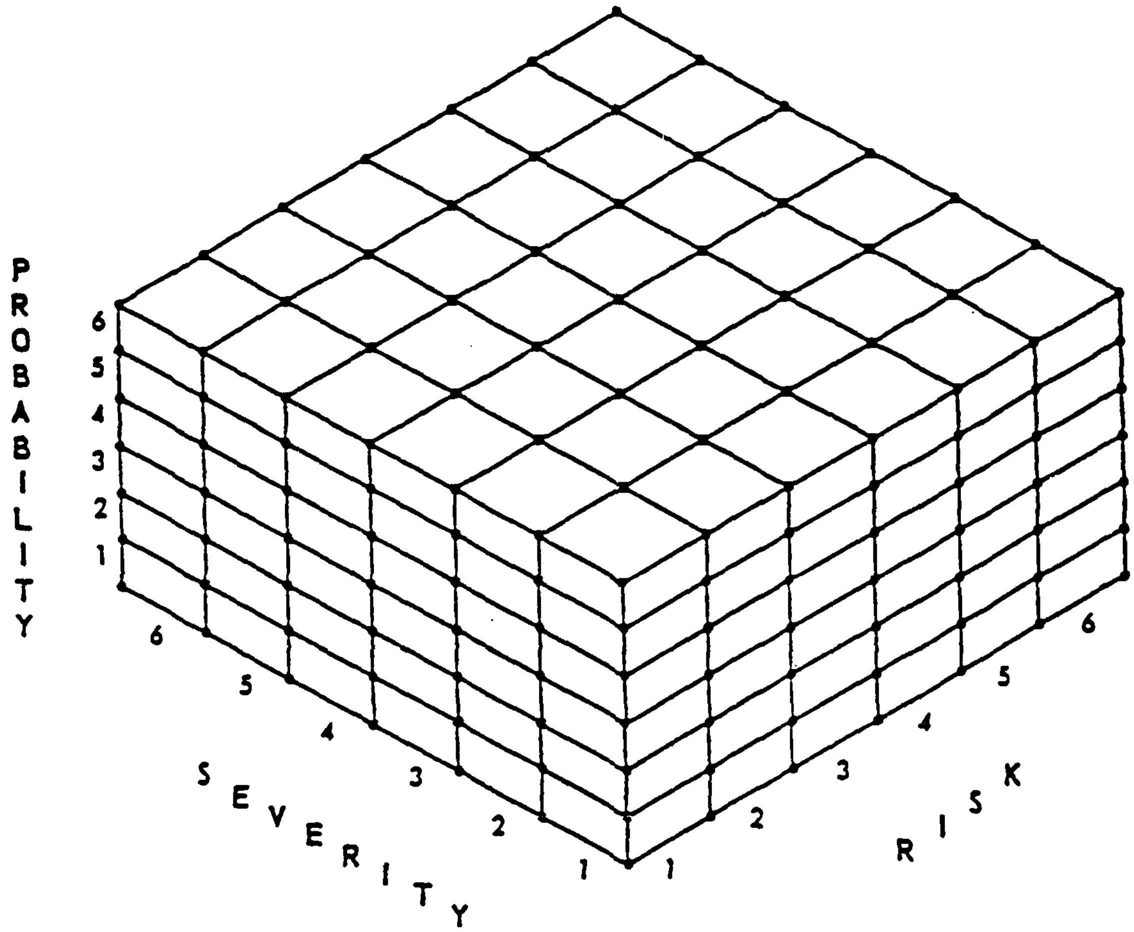


Figure 15 Probability-Risk-Severity Relationship



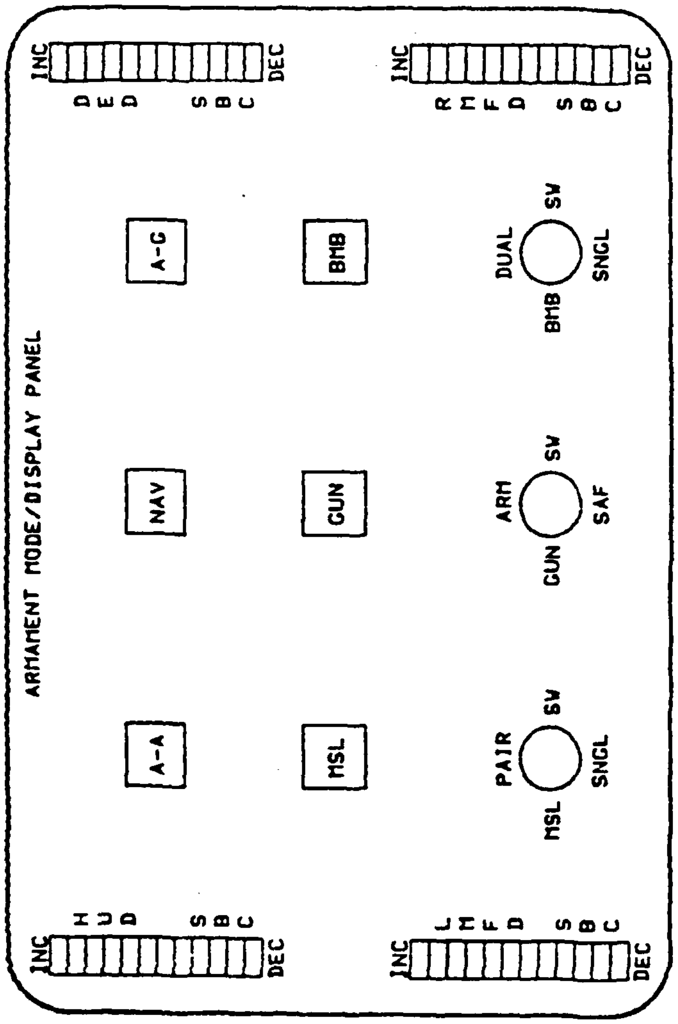
Probability-Risk-Severity Relationship

COMPUTERIZED TASK ANALYSIS DATA BASE BUILDING BLOCKS

1. GLOSSARY FILE
2. TASK EVENT FILE
3. TASK TIME FILE
4. TASK LINKING
5. TASK SIMULATION

GLOSSARY FILE

* NOMENCLATURE
* DEFINITIONS
* MNEMONICS
* PUSHBUTTONS
* CONTROLS
* SWITCHES
* KNOBS


ARMAMENT MODE/DISPLAY PANEL
INC
DEC
HUD
SBC
DED
SBC
LMFD
SBC
RMFD
SBC
A-A
NAV
A-G
MSL
GUN
BMB
MSL
PAIR
SV
SNGL
GUN
ARM
SV
SAF
BMB
DUAL
SV
SNGL

```
NOMENCLATURE         [~N]
ON                   [/ ]
ARMAMENT MODE PANEL [AMD]
--------------------------

A-A>~N/AMD
A-G~N/AMD
ARM~N/AMD
BMB>~N/AMD
BMB SW>~N/AMD
  .
  .
  .
```

NOMENCLATURE [¯N]
ON [/]
ARMAMENT MODE/DISPLAY PANEL [AMD]

A-A>¯N/AMD
A-G>¯N/AMD
ARM>¯N/AMD
BMB>¯N/AMD
BMB SW>¯N/AMD
BMB SW-DUAL¯N/AMD
BMB SW-SNGL¯N/AMD
DEC>¯N/AMD
DED>¯N/AMD
DED SBC>¯N/AMD
DED SBC-DEC>¯N/AMD
DED SBC-INC>¯N/AMD
DUAL>¯N/AMD
GUN>¯N/AMD
GUN SW>¯N/AMD
GUN SW-ARM>¯N/AMD
GUN SW-SAF>¯N/AMD
HUD>¯N/AMD
HUD SBC>¯N/AMD
HUD SBC-DEC>¯N/AMD
HUD SBC-INC>¯N/AMD
INC>¯N/AMD
LMFD>¯N/AMD
LMFD SBC>¯N/AMD
LMFD SBC-DEC>¯N/AMD
LMFD SBC-INC>¯N/AMD
MSL>¯N/AMD
MSL SW>¯N/AMD
MSL SW-PAIR>¯N/AMD
MSL SW-SNGL>¯N/AMD
NAV>¯N/AMD
PAIR>¯N/AMD
RMFD>¯N/AMD
RMFD SBC>¯N/AMD
RMFD SBC-DEC>¯N/AMD
RMFD SBC-INC>¯N/AMD
SAF>¯N/AMD
SBC>¯N/AMD
SNGL>¯N/AMD
SW>¯N/AMD

```
NOMENCLATURE DEFINITIONS [<---------->]
MNEMONIC                 [~M/]
PUSHBUTTON               [~P/]
CONTROL                  [~C/]
SWITCH                   [~S/]
KNOB                     [~K/]
--------------------------------


A-A<A-A MODE>~M/HUD/MFD~N/AMD~P/AMD
A-G<A-G MODE>~M/HUD/MFD~N/AMD~P/AMD
ARM~M/HUD/MFD~N/AMD
BMB<BMB MODE>~M/HUD/MFD~N/AMD~P/AMD
BMB SW>~N/AMD~C/AMD
```

TASK EVENT FILE

* TASK NAME
  (FOUR LETTER CODE)
* TASK DEFINITION
  (TEXT/ABBREVIATIONS)

NOMENCLATURE DEFINITIONS [<---------->]
MNEMONIC [¯M/]
PUSHBUTTON [¯P/]
CONTROL [¯C/]
SWITCH [¯S/]
KNOB [¯K/]

A-A<A-A MODE>¯M/HUD/MFD¯N/AMD¯P/AMD
A-G<A-G MODE>¯M/HUD/MFD¯N/AMD¯P/AMD
ARM<FIRING MODE>¯M/HUD/MFD¯N/AMD
BMB<BMB MODE>¯M/HUD/MFD¯N/AMD¯P/AMD
BMB SW<BMB DUAL/SNGL SWITCH>¯N/AMD¯S/AMD
BMB SW-DUAL<DOUBLE BOMB RELEASE>¯N/AMD¯S/AMD
BMB SW-SNGL<SINGLE BOMB RELEASE>¯N/AMD¯S/AMD
DEC<DECREASE INTENSITY>¯N/AMD
DED<DATA ENTRY DISPLAY>¯N/AMD
DED SBC<DED SBC KNOB>¯N/AMD¯C/AMD
DED SBC-DEC<SYMBOLOGY DECREASE INTENSITY>¯N/AMD¯K/AMD
DED SBC-INC<SYMBOLOGY INCREASE INTENSITY>¯N/AMD¯K/AMD
DUAL<DOUBLE BOMB RELEASE>¯M/HUD/MFD¯N/AMD
GUN<GUN MODE>¯M/HUD/MFD¯N/AMD
GUN SW<GUN ARM/SAF SWITCH>¯N/AMD¯C/AMD
GUN SW-ARM<GUN IN ARM MODE>¯N/AMD¯S/AMD
GUN SW-SAF<GUN IN SAFE MODE>¯N/AMD¯S/AMD
HUD<HEAD-UP DISPLAY>¯N/AMD
HUD SBC<HUD SBC KNOB>¯N/AMD¯C/AMD
HUD SBC-DEC<SYMBOLOGY DECREASE INTENSITY>¯N/AMD¯K/AMD
HUD SBC-INC<SYMBOLOGY INCREASE INTENSITY>¯N/AMD¯K/AMD
INC<INCREASE INTENSITY>¯N/AMD
LMFD<LEFT MULTIFUNCTION DISPLAY>¯N/AMD
LMFD SBC<LMFD SBC KNOB>¯N/AMD¯C/AMD
LMFD SBC-DEC<SYMBOLOGY DECREASE INTENSITY>¯N/AMD¯K/AMD
LMFD SBC-INC<SYMBOLOGY INCREASE INTENSITY>¯N/AMD¯K/AMD
MSL<MISSILE MODE>¯N/AMD¯P/AMD
MSL SW<MISSILE PAIR/SNGL SWITCH>¯N/AMD¯C/AMD
MSL SW-PAIR<MISSILE FIRING IN PAIR>¯N/AMD¯S/AMD
MSL SW-SNGL<SINGLE MISSILE FIRING>¯N/AMD¯S/AMD
NAV<NAVIGATION MODE>¯M/HUD/MFD¯N/AMD¯P/AMD
PAIR<DOUBLE FIRING>¯M/HUD/MFD¯N/AMD
RMFD<RIGHT MULTIFUNCTION DISPLAYU>¯N/AMD
RMFD SBC<RMFD SBC KNOB>¯N/AMD¯C/AMD
RMFD SBC-DEC<SYMBOLOGY DECREASE INTENSITY>¯N/AMD¯K/AMD
RMFD SBC-INC<SYMBOLOGY INCREASE INTENSITY>¯N/AMD¯K/AMD
SAF<NON-FIRING MODE>¯M/HUD/MFD¯N/AMD
SBC<SYMBOLOGY BRIGHTNESS/CONTRAST>¯N/AMD
SNGL<SINGLE BMB RELEASE/SINGLE MSL FIRING>¯M/HUD/MFD¯N/AMD
SW<SWITCH>¯N/AMD

TASK TIME FILE

* MINIMUM TASK TIME
  (ESTIMATED OR OBSERVED)
* MAXIMUM TASK TIME
  (ESTIMATED OR OBSERVED)
* EXPECTED TASK TIME
  (COMPUTATION - MONTE CARLO)

```
HA-A-OBSERVE HUD TO VERIFY A-A MNEMONIC
HA-A  0.5/   3.5/   0.0/  0
MA-A-OBSERVE MFD TO VERIFY A-A MNEMONIC
MA-A  0.5/   3.5/   0.0/  0
PA-A-DEPRESS/RELEASE A-A PUSHBUTTON ON AMD
PA-A  1.5/   4.0/   0.0/   0
```

```
ABM1-MOVE BMB SW TO DOUBLE BOMB RELEASE POSITION
ABM1  2.5/  5.0/  0.0/  0
ABM2-MOVE BMB SW TO SINGLE BOMB RELEASE POSITION
ABM2  2.5/  5.0/  0.0/  0
ABMB-POSITION BMB DUAL/SNGL SW ON AMD AS DESIRED
ABMB  2.5/  5.0/  0.0/  2 0 ABM1ABM2
ADE1-ROTATE DED SBC KNOB TO SYMBOLOGY DECREASE INTENSITY POSITION
ADE1  2.5/  5.0/  0.0/  0
ADE2-ROTATE DED SBC KNOB TO SYMBOLOGY INCREASE INTENSITY POSITION
ADE2  2.5/  5.0/  0.0/  0
ADED-POSITION DED SBC KNOB ON AMD AS DESIRED
ADED  2.5/  5.0/  0.0/  2 0 ADE1ADE2
AGU1-MOVE GUN SW TO GUN IN ARM MODE POSITION
AGU1  2.5/  5.0/  0.0/  0
AGU2-MOVE GUN SW TO GUN IN SAFE MODE POSITION
AGU2  2.5/  5.0/  0.0/  0
AGUN-POSITION GUN ARM/SAF SWITCH ON AMD AS DESIRED
AGUN  2.5/  5.0/  0.0/  2 0 AGU1AGU2
AHU1-ROTATE HUD SBC KNOB TO SYMBOLOGY DECREASE INTENSITY POSITION
AHU1  2.5/  5.0/  0.0/  0
AHU2-ROTATE HUD SBC KNOB TO SYMBOLOGY INCREASE INTENSITY POSITION
AHU2  2.5/  5.0/  0.0/  0
AHUD-POSITION HUD SBC KNOB ON AMD AS DESIRED
AHUD  2.5/  5.0/  0.0/  2 0 AHU1AHU2
ALM1-ROTATE LMFD SBC KNOB TO SYMBOLOGY DECREASE INTENSITY POSITION
ALM1  2.5/  5.0/  0.0/  0
ALM2-ROTATE LMFD SBA KNOB TO SYMBOLOGY INCREASE INTENSITY POSITION
ALM2  2.5/  5.0/  0.0/  0
ALMF-POSITION LMFD SBC KNOB ON AMD AS DESIRED
ALMF  2.5/  5.0/  0.0/  2 0 ALM1ALM2
AMS1-MOVE MSL SW TO MISSILE FIRING IN PAIR POSITION
AMS1  2.5/  5.0/  0.0/  0
AMS2-MOVE MSL SW TO SINGLE MISSILE FIRING POSITION
AMS2  2.5/  5.0/  0.0/  0
AMSL-POSITION MISSILE PAIR/SNGL SWITCH ON AMD AS DESIRED
AMSL  2.5/  5.0/  0.0/  2 0 AMS1AMS2
ARM1-ROTATE RMFD SBC KNOB TO SYMBOLOGY INCREASE INTENSITY POSITION
ARM1  2.5/  5.0/  0.0/  0
ARM2-ROTATE RMFD SBC KNOB TO SYMBOLOGY DECREASE INTENSITY POSITION
ARM2  2.5/  5.0/  0.0/  0
ARMF-POSITION RMFD SBC KNOB ON AMD AS DESIRED
ARMF  2.5/  5.0/  0.0/  2 0 ARM1ARM2
HA-A-OBSERVE HUD TO VERIFY A-A MNEMONIC
HA-A  0.5/  3.5/  0.0/  0
HA-G-OBSERVE HUD TO VERIFY A-G MNEMONIC
HA-G  0.5/  3.5/  0.0/  0
HARM-OBSERVE HUD TO VERIFY ARM MNEMONIC
HARM  0.5/  3.5/  0.0/  0
HBMB-OBSERVE HUD TO VERIFY BMB MNEMONIC
HBMB  0.5/  3.5/  0.0/  0
HDUA-OBSERVE HUD TO VERIFY DUAL MNEMONIC
HDUA  0.5/  3.5/  0.0/  0
HGUN-OBSERVE HUD TO VERIFY GUN MNEMONIC
HGUN  0.5/  3.5/  0.0/  0
HMSL-OBSERVE HUD TO VERIFY MSL MNEMONIC
HMSL  0.5/  3.5/  0.0/  0
HNAV-OBSERVE HUD TO VERIFY NAV MNEMONIC
HNAV  0.5/  3.5/  0.0/  0
HPAI-OBSERVE HUD TO VERIFY PAIR MNEMONIC
```

```
HPAI  0.5/  3.5/  0.0/  0
HSAF-OBSERVE HUD TO VERIFY SAF MNEMONIC
HSAF  0.5/  3.5/  0.0/  0
HSNG-OBSERVE HUD TO VERIFY SNGL MNEMONIC
HSNG  0.5/  3.5/  0.0/  0
MA-A-OBSERVE MFD TO VERIFY A-A MNEMONIC
MA-A  0.5/  3.5/  0.0/  0
MA-G-OBSERVE MFD TO VERIFY A-G MNEMONIC
MA-G  0.5/  3.5/  0.0/  0
MARM-OBSERVE MFD TO VERIFY ARM MNEMONIC
MARM  0.5/  3.5/  0.0/  0
MBMB-OBSERVE MFD TO VERIFY BMB MNEMONIC
MBMB  0.5/  3.5/  0.0/  0
MDUA-OBSERVE MFD TO VERIFY DUAL MNEMONIC
MDUA  0.5/  3.5/  0.0/  0
MGUN-OBSERVE MFD TO VERIFY GUN MNEMONIC
MGUN  0.5/  3.5/  0.0/  0
MMSL-OBSERVE MFD TO VERIFY MSL MNEMONIC
MMSL  0.5/  3.5/  0.0/  0
MNAV-OBSERVE MFD TO VERIFY NAV MNEMONIC
MNAV  0.5/  3.5/  0.0/  0
MPAI-OBSERVE MFD TO VERIFY PAIR MNEMONIC
MPAI  0.5/  3.5/  0.0/  0
MSAF-OBSERVE MFD TO VERIFY SAF MNEMONIC
MSAF  0.5/  3.5/  0.0/  0
MSNG-OBSERVE MFD TO VERIFY SNGL MNEMONIC
MSNG  0.5/  3.5/  0.0/  0
PA-A-DEPRESS/RELEASE A-A PUSHBUTTON ON AMD
PA-A  1.5/  4.0/  0.0/  0
PA-G-DEPRESS/RELEASE A-G PUSHBUTTON ON AMD
PA-G  1.5/  4.0/  0.0/  0
PBMB-DEPRESS/RELEASE BMB PUSHBUTTON ON AMD
PBMB  1.5/  4.0/  0.0/  0
PGUN-DEPRESS/RELEASE GUN PUSHBUTTON ON AMD
PGUN  1.5/  4.0/  0.0/  0
PMSL-DEPRESS/RELEASE MSL PUSHBUTTON ON AMD
PMSL  1.5/  4.0/  0.0/  0
PNAV-DEPRESS/RELEASE NAV PUSHBUTTON ON AMD
PNAV  1.5/  4.0/  0.0/  0
```

TASK LINKING

-AND- GATES

-OR- GATES

TASK EVENT CHAINS

```
   50
ABM1  2.5/  5.0/  3.1/  0
ABM2  2.5/  5.0/  3.2/  0
ABMB  2.5/  5.0/  3.2/  2  0  ABM1ABM2
ADE1  2.5/  5.0/  3.2/  0
ADE2  2.5/  5.0/  3.3/  0
ADED  2.5/  5.0/  3.3/  2  0  ADE1ADE2
AGU1  2.5/  5.0/  3.3/  0
AGU2  2.5/  5.0/  3.3/  0
AGUN  2.5/  5.0/  3.3/  2  0  AGU1AGU2
AHU1  2.5/  5.0/  3.4/  0
AHU2  2.5/  5.0/  3.4/  0
AHUD  2.5/  5.0/  3.4/  2  0  AHU1AHU2
ALM1  2.5/  5.0/  3.4/  0
ALM2  2.5/  5.0/  3.4/  0
ALMF  2.5/  5.0/  3.4/  2  0  ALM1ALM2
AMS1  2.5/  5.0/  3.4/  0
AMS2  2.5/  5.0/  3.5/  0
AMSL  2.5/  5.0/  3.5/  2  0  AMS1AMS2
ARM1  2.5/  5.0/  3.5/  0
ARM2  2.5/  5.0/  3.5/  0
ARMF  2.5/  5.0/  3.5/  2  0  ARM1ARM2
HA-A  0.5/  3.5/  2.6/  0
HA-G  0.5/  3.5/  2.5/  0
HARM  0.5/  3.5/  2.5/  0
HBMB  0.5/  3.5/  2.5/  0
HDUA  0.5/  3.5/  2.4/  0
HGUN  0.5/  3.5/  2.4/  0
HMSL  0.5/  3.5/  2.4/  0
HNAV  0.5/  3.5/  2.4/  0
HPAI  0.5/  3.5/  2.4/  0
HSAF  0.5/  3.5/  2.4/  0
HSNG  0.5/  3.5/  2.4/  0
MA-A  0.5/  3.5/  2.3/  0
MA-G  0.5/  3.5/  2.3/  0
MARM  0.5/  3.5/  2.3/  0
MBMB  0.5/  3.5/  2.3/  0
MDUA  0.5/  3.5/  2.3/  0
MGUN  0.5/  3.5/  2.3/  0
MMSL  0.5/  3.5/  2.3/  0
MNAV  0.5/  3.5/  2.3/  0
MPAI  0.5/  3.5/  2.2/  0
MSAF  0.5/  3.5/  2.2/  0
MSNG  0.5/  3.5/  2.2/  0
PA-A  1.5/  4.0/  2.2/  0
PA-G  1.5/  4.0/  2.3/  0
PBMB  1.5/  4.0/  2.3/  0
PGUN  1.5/  4.0/  2.3/  0
PMSL  1.5/  4.0/  2.3/  0
PNAV  1.5/  4.0/  2.3/  0
```

```
PHASE:  VA-G       WEAPON:  65AGM      MINIMUM PHASE TIME:     9.7    MAXIMUM PHASE TIME:     9.7
SPEED:   600       SOI:     HUD                  DISTANCE:     9829             DISTANCE:     9829
WTS:     3.2       PRD:      0.0                 RUN TIME:     9.7              RUN TIME:     9.7
DELTA MIN:    0.0  MAX:      0.0                 DISTANCE:     9329             DISTANCE:     9829
EXPECTED DISTANCE:    11000        DISTANCE REMAINING:        34171   DISTANCE REMAINING:    38171
```

VA-G
9.7

A

A-GM	HA-G	RA-G	LA-G
2.0	2.6	2.5	2.6

VA-G-VERIFY A-G MODE SELECTED
A-GM-DEPR/REL A-G PB ON AMDP TO ENGAGE A-G MODE
HA-G-OBSERVE HUD TO VERIFY A-G MNEMONIC & A-G MODE SYMBOLOGY
RA-G-OBSERVE RMFD TO VERIFY A-G MNEMONIC & A-G MODE SYMBOLOGY
LA-G-OBSERVE LMFD TO VERIFY A-G MNEMONIC & A-G MODE SYMBOLOGY

```
---------------------------------------------------------------------------------------------------
PHASE:  VBMB        WEAPON:  GBU        MINIMUM PHASE TIME:      9.0     MAXIMUM PHASE TIME:      9.0
SPEED:   600        SOI:     FCR                  DISTANCE:     9120               DISTANCE:     9120
WTS:     2.5        PRD:      0.0                 RUN TIME:     18.7               RUN TIME:     18.7
DELTA MIN:    0.0   MAX:      0.0                 DISTANCE:    18949               DISTANCE:    18949
EXPECTED DISTANCE:     12000            DISTANCE REMAINING:    25051     DISTANCE REMAINING:    29051
---------------------------------------------------------------------------------------------------
----
VBMB
 9.0
----
  A
---- ---- ---- ----
BMBM HBMB RBMB LBMB
 2.0  2.0  2.5  2.5
---- ---- ---- ----

---------------------------------------------------------------------------------------------------
VBMB-VERIFY BMB MODE SELECTED
BMBM-DEPR/REL BMB PB ON AMDP TO ENGAGE BOMB MODE
HBMB-OBSERVE HUD TO VERIFY BMB MNEMONIC & BMB MODE SYMBOLOGY
RBMB-OBSERVE RMFD TO VERIFY BMB MNEMONIC & BMB SYMBOLOGY
LBMB-OBSERVE LMFD TO VERIFY BMB MNEMONIC & BMB SYMBOLOGY
```

```
PHASE:  VA-A        WEAPON:  51GUN      MINIMUM PHASE TIME:     9.7     MAXIMUM PHASE TIME:     9.7
SPEED:   650        SOI:       HUD                DISTANCE:   10648               DISTANCE:   10648
WTS:     4.1        PRD:       0.0                RUN TIME:    28.4               RUN TIME:    28.4
DELTA MIN:     0.0   MAX:      0.0                DISTANCE:   29597               DISTANCE:   29597
EXPECTED DISTANCE:     12000             DISTANCE REMAINING:   14403     DISTANCE REMAINING:   18403
```

VA-A
9.7

A

A-AM	HA-A	RA-A	LA-A
1.9	2.6	2.6	2.6

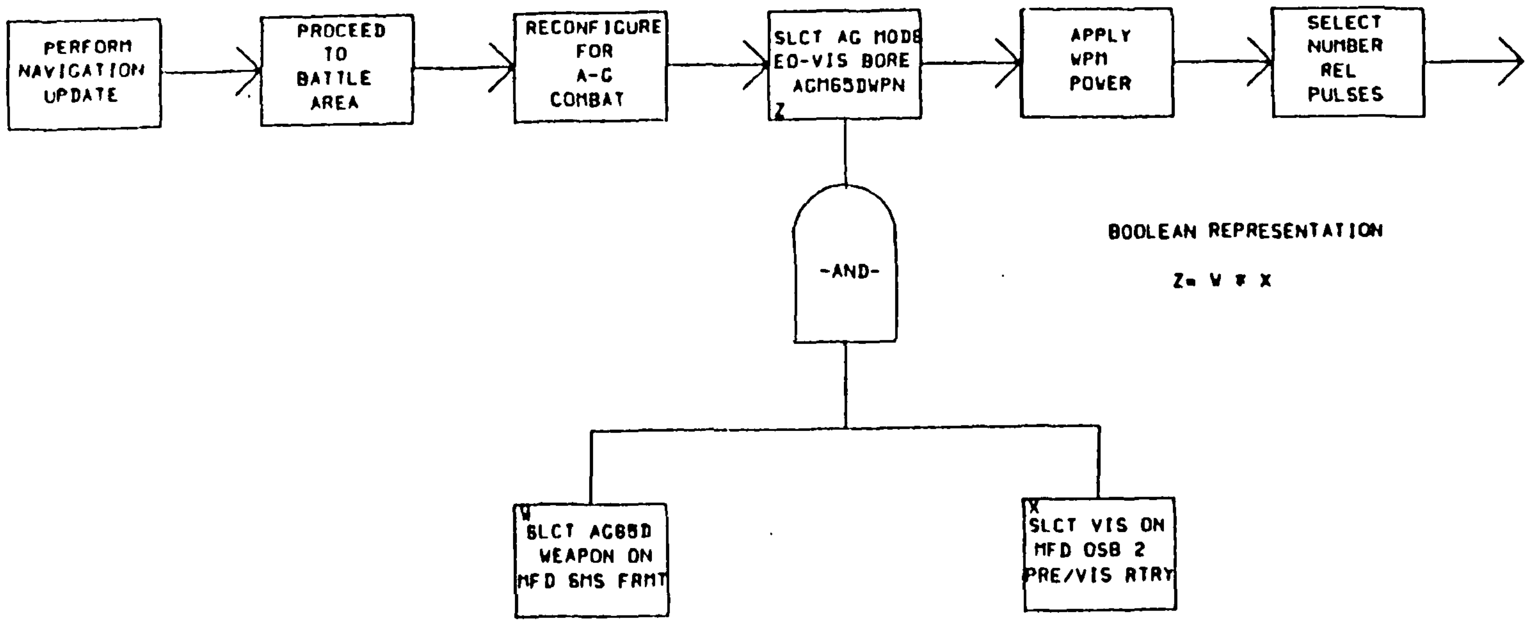
VA-A-VERIFY A-A MODE SELECTED
A-AM-DEPR/REL A-A PB ON AMDP TO ENGAGE A-A MODE
HA-A-OBSERVE HUD TO VERIFY A-A MNEMONIC & A-A MODE SYMBOLOGY
RA-A-OBSERVE RMFD TO VERIFY A-A MNEMONIC & A-A MODE SYMBOLOGY
LA-A-OBSERVE LMFD TO VERIFY A-A MNEMONIC & A-A MODE SYMBOLOGY

```
--------------------------------------------------------------------------------------------------------------
PHASE:  VGUN      WEAPON:  51GUN      MINIMUM PHASE TIME:      9.9      MAXIMUM PHASE TIME:      9.
SPEED:   650      SOI:     HUD                  DISTANCE:    10868                DISTANCE:   1086
WTS:     4.6      PRD:      0.0                 RUN TIME:     38.3                RUN TIME:    38.
DELTA MIN:    0.0   MAX:    0.0                 DISTANCE:    40465                DISTANCE:   4046
EXPECTED DISTANCE:    13000           DISTANCE REMAINING:     3535      DISTANCE REMAINING:    753
--------------------------------------------------------------------------------------------------------------
----
VGUN
 9.9
----
  A
---- ---- ---- ----
GUNM HGUN RGUN LGUN
 2.9  2.0  2.5  2.5
---- ---- ---- ----

--------------------------------------------------------------------------------------------------------------
VGUN-VERIFY GUN MODE SELECTED
GUNM-DEPR/REL GUN PB ON AMDP TO ENGAGE GUN MODE
HGUN-OBSERVE HUD TO VERIFY GUN MNEMONIC & GUN MODE SYMBOLOGY
RGUN-OBSERVE RMFD TO VERIFY GUN MNEMONIC & GUN SYMBOLOGY
LGUN-OBSERVE LMFD TO VERIFY GUN MNEMONIC & GUN SYMBOLOGY
```

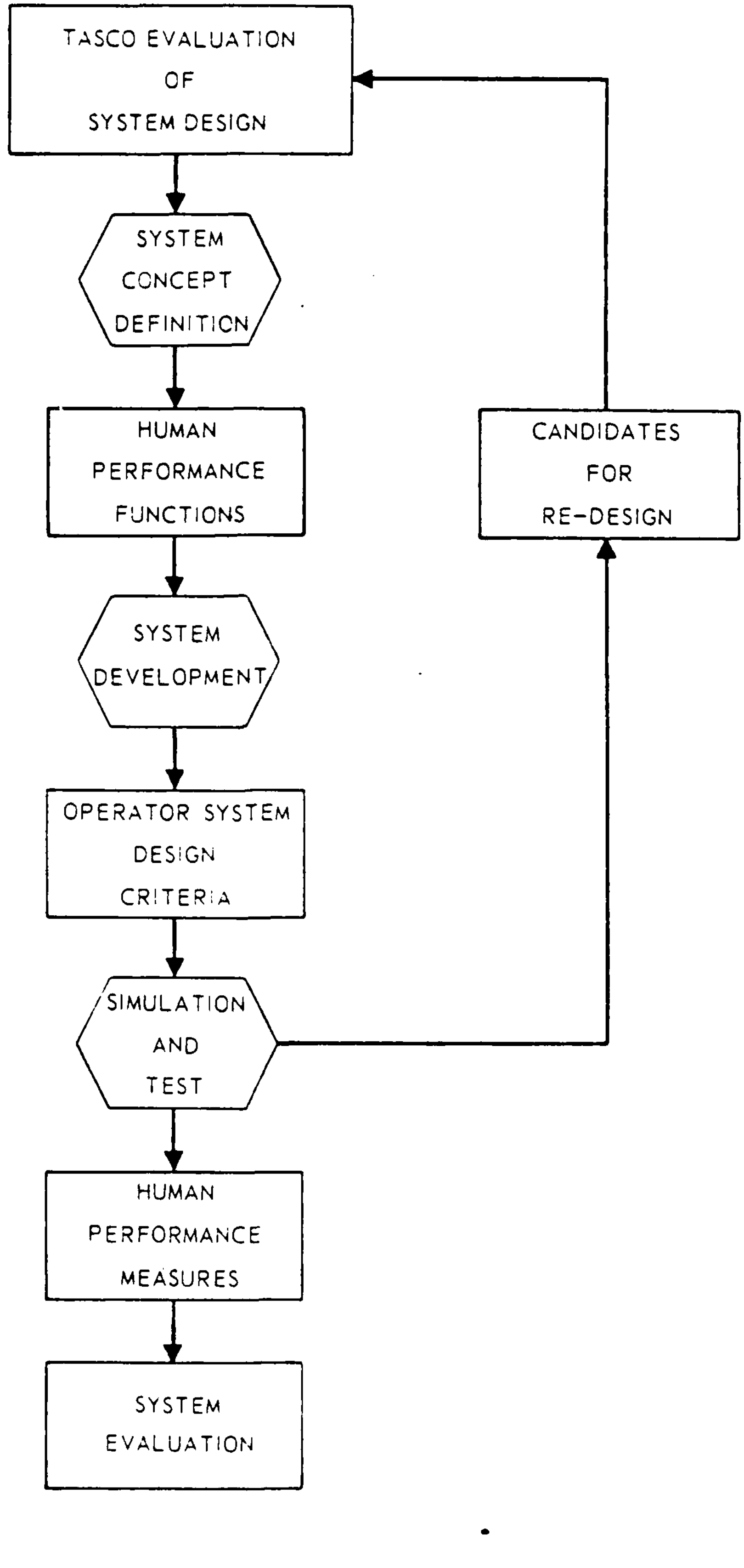
SYSTEMS APPROACH CRITICAL PILOT-AVIONIC TASK ANALYSIS
[SACPATA]

Mission Task Flow
Basic Element Task Analysis
(Beta Factors)



Task Element
Absolute Logic Probability Hazard Analysis
(Alpha Factors)

**** HIT <RETURN> TO CONTINUE ****


TASCO EVALUATION
OF
SYSTEM DESIGN
SYSTEM
CONCEPT
DEFINITION
HUMAN
PERFORMANCE
FUNCTIONS
SYSTEM
DEVELOPMENT
OPERATOR SYSTEM
DESIGN
CRITERIA
SIMULATION
AND
TEST
HUMAN
PERFORMANCE
MEASURES
SYSTEM
EVALUATION
CANDIDATES
FOR
RE-DESIGN

Procedure for accessing computers through the EDWARDS TAC.

1. Make sure your terminal and modem are on and happy.

2. Call one of the TAC numbers listed below:

   277-7949
   277-7966
   277-7968
   277-7970
   277-7971
   277-7972
   277-7973
   277-7974

3. Once the TAC answers and your modem is connected with it, type a "CONTROL-Q". You do this by holding down the key marked "CTRL" and simultaneously hitting the "Q" key.

4. The TAC will then respond with a sign-on message, and await your command.

5. Depending on the computer you want to connect to, type:

```
@o 26.1.0.39          (for the Edwards-2060)
@o 26.0.0.39          (for the Edwards-VAX)
@o 67                 (for the AFSC-HQ at Andrews)
```

6. The TAC will then prompt for USER ID and ACCESS CODE. Enter these as they are written on your TAC ACCESS CARD.

7. You will then see some "progress of connection" messages.

8. Hit <RETURN> a couple times to get the destination computer's attention.

9. Once the destination computer responds, you may log in just as if you were at the machine's site.

****

10. When you are done on the destination computer, type "LOGO" to logout. You will be returned to the Edwards TAC, where you may point at another destination computer, or simply type "@l" to log out of the TAC.

11. Hang up, and you're done...

SESSION 3. PHYSIOLOGICAL TECHNIQUES

Presenters

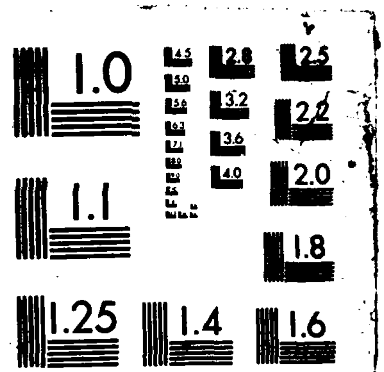
G. F. Wilson: Neuropsychological Workload Test Battery

J.A. Stern: Workload - A Psychophysiological Approach

AD-A185 650 PROCEEDINGS OF THE DOD WORKLOAD ASSESSMENT WORKSHOP ON WORKLOAD ASSESSMEN..(U) NAVAL UNDERWATER SYSTEMS CENTER NEWPORT RI H M FIEDLER 15 SEP 87 NUSC-TD-6608

4/4

UNCLASSIFIED F/G 23/2 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

NEUROPSYCHOLOGICAL WORKLOAD

HARRY G. ARMSTRONG

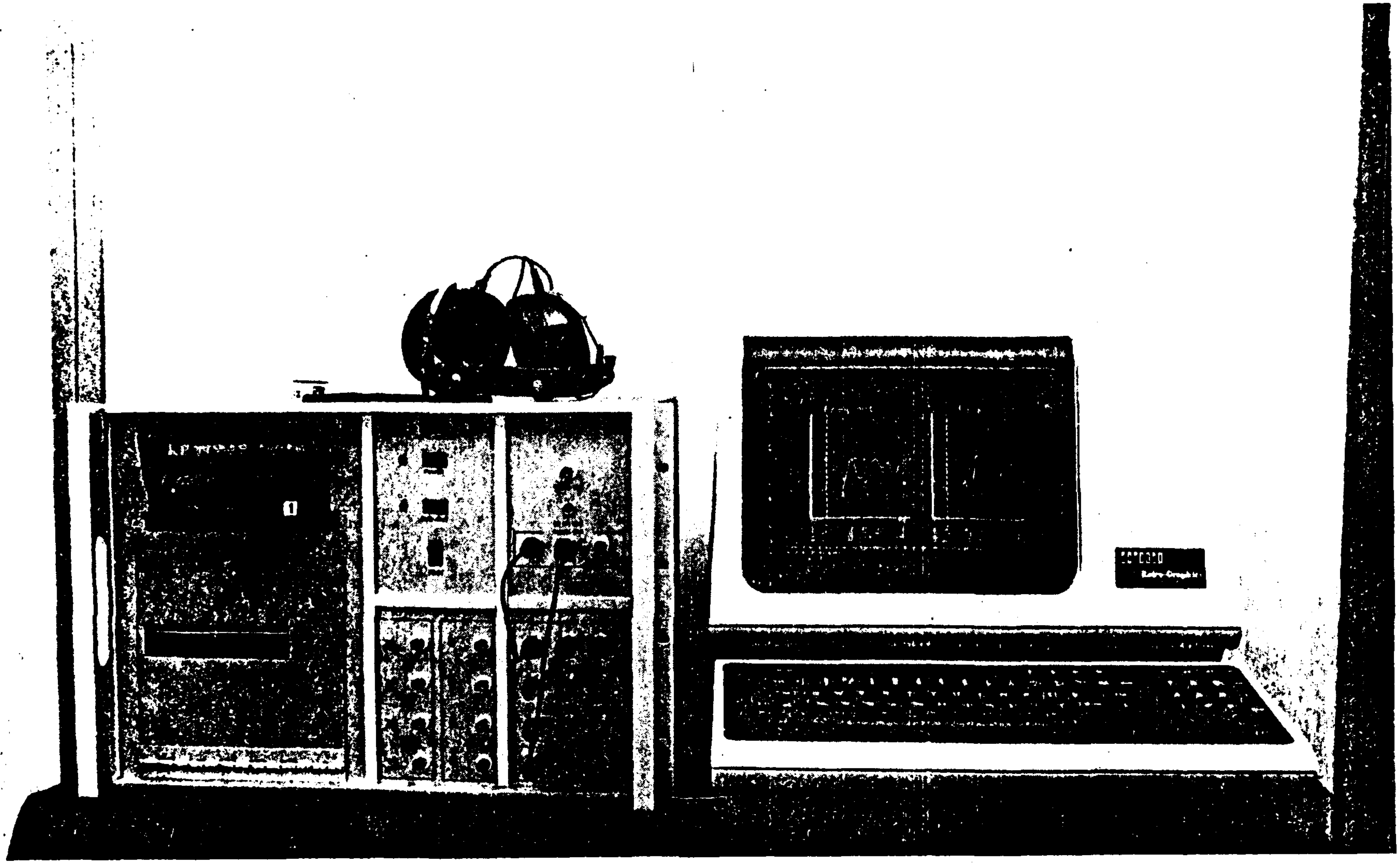
AEROSPACE MEDICAL RESEARCH LABORATORY

HUMAN ENGINEERING DIVISION

WORKLOAD AND ERGONOMICS BRANCH

WRIGHT-PATTERSON AIR FORCE BASE, OHIO

GLENN F. WILSON, Ph.D.
AAMRL/HEG
WRIGHT-PATTERSON, OH
45433-6573
513-255-8748



NEUROPSYCHOLOGICAL WORKLOAD TEST BATTERY
NWTB

ARMSTRONG AEROSPACE MEDICAL RESEARCH LABORATORY

NEUROPSYCHOLOGICAL WORKLOAD TEST BATTERY (NWTB)

CENTRAL MEASURES

- TRANSIENT EVOKED RESPONSE
  - — ODD-BALL
  - — MEMORY SCAN
  - — MEMORY UPDATE
  - — SELECTIVE ATTENTION
  - — MONITORING
  - — BRAIN STEM
  - — TRACKING
  - — FLASH
- STEADY STATE EVOKED RESPONSE
  - — CHECKER BOARD
  - — SINE WAVE GRATING
  - — UNPATTERNED FIELD

PERIPHERAL MEASURES

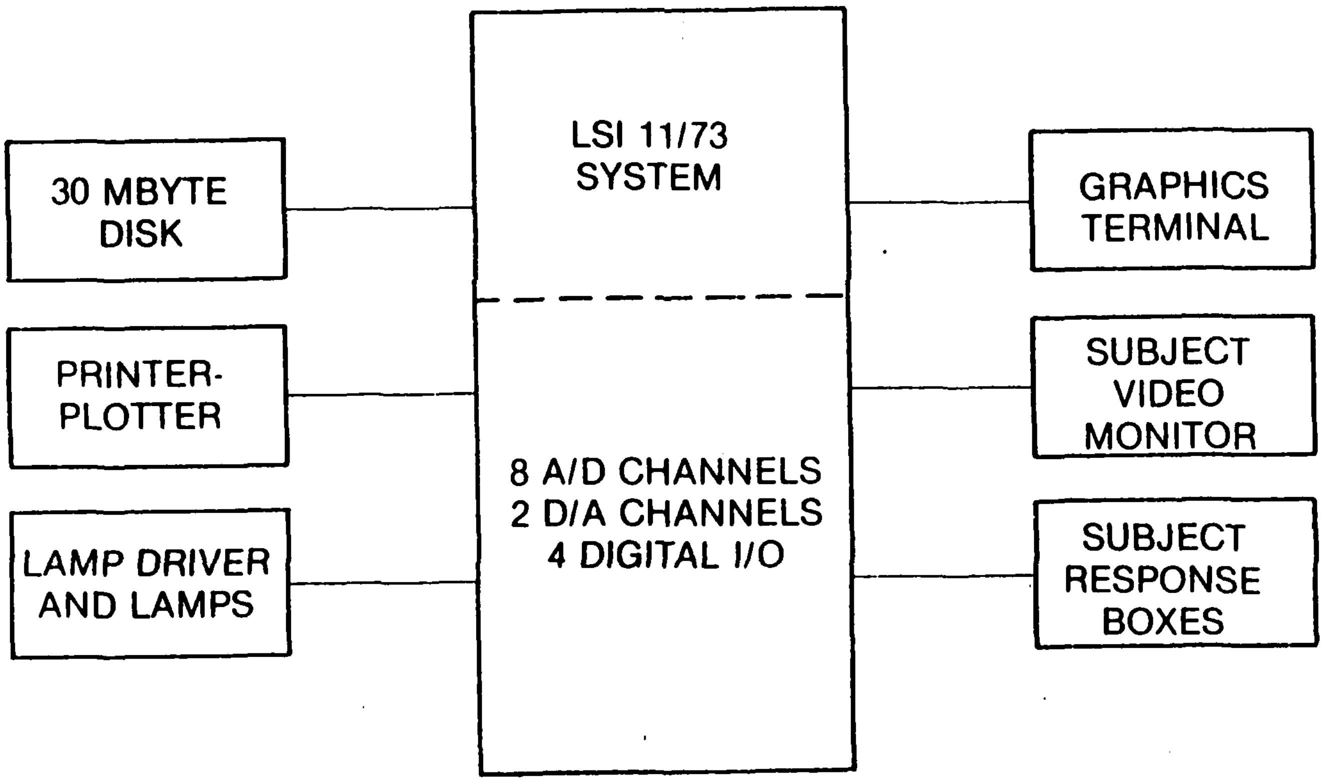
- EYE BLINK
- HEART RATE
- MUSCLE

PERFORMANCE MEASURES

- REACTION TIME
- ERROR SCORES

HE-85-5-32

COMPONENTS OF THE NWTB



HE-85-9-92

TEST BATTERY MENU

FILE IDENTIFICATION DLO: XXXXXX.XXX

1. AUDIO RARE EVENT MONITORING
2. VISUAL RARE EVENT MONITORING
3. HEARTRATE
4. STEADY STATE EVOKED RESPONSE
5. HIGH FREQUENCY STEADY STATE
6. EMG ANALYSIS
7. STERNBERG TEST
8. BRAIN STEM TEST
9. EYEBLINK TEST
11. CRITICAL TRACKING TEST
12. CONTINUOUS PERFORMANCE
13. SELECTIVE ATTENTION
16. START TESTS
17. CHANGE FILE ID

ENTER NUMBER OF SELECTED REQUEST

HE 86 9 118

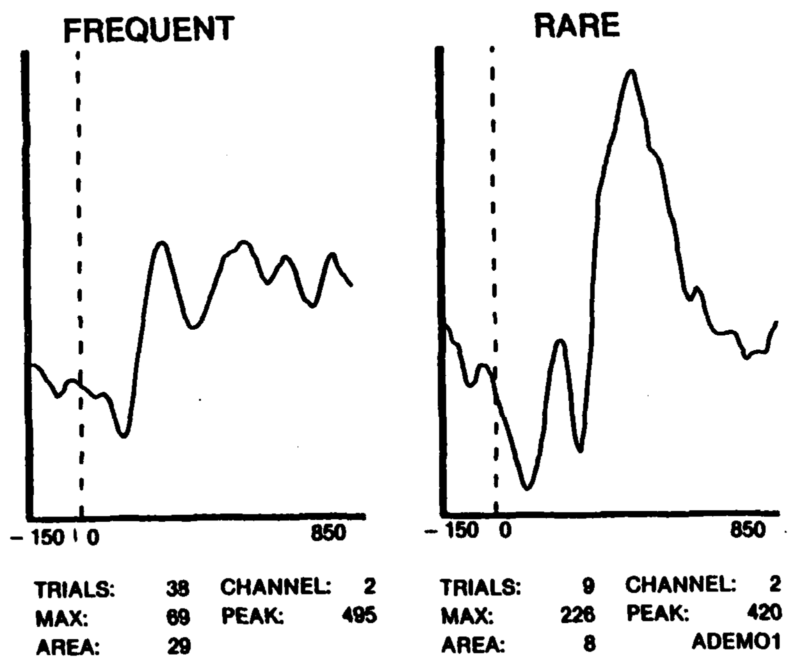
AUDIO RARE EVENT MONITORING

AUDIO RARE EVENT MONITORING

1	FREQ OF RARE TONE (500-2000 Hz)	1500
2	FREQ OF COMMON TONE (500-2000 Hz)	1200
3	INTERSTIMULUS INTERVAL (SEC)	2
4	PROB OF RARE EVENT (10,20,30,40)	20
5	FIXED PROB FOR RARE TONE (Y/N)	Y
6	DURATION OF TEST (1-30 MIN)	3
7	REQ NO OF RARE EVENTS (0-100)	5
8	REJECTION THRESHOLD (EOG 0-2048)	262
9	REJECTION THRESHOLD (EMG 0-2048)	2048
10	TONE INTENSITY (1-10)	5
11	EXTERNAL TRIGGER (Y/N)	N
12	SAVE RAW DATA? (Y/N)	Y

21 END
22 ABORT

ENTER PARAMETER NO, SPACE, AND NEW PARAMETER



HEART RATE ANAYSIS

HEART RATE

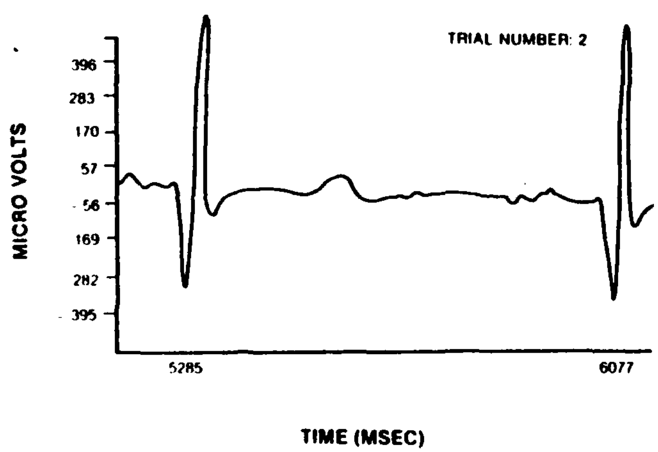
1 NUMBER OF TRIALS (1-100)	3
2 INTERVAL BETWEEN TRIALS (0-60 SEC) (0 = CONTINUOUS)	0
LENGTH OF EACH TRIAL (6-60 SEC) (INCREMENTS OF 6 SEC)	30
4 EXTERNAL TRIGGER (Y/N)	N
5 SAVE RAW DATA? (Y/N)	Y

21 END
22 ABORT

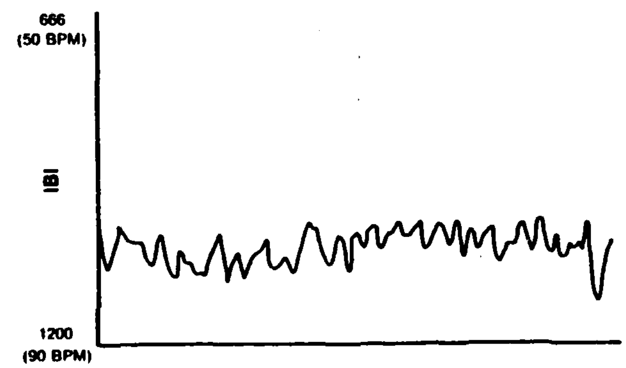
HEART RATE ANALYSIS

CARDIAC AMPLITUDE:	30
DIFFERENCE CRITERION:	199
MAXIMUM TIME BETWEEN R WAVES:	1000
MINIMUM TIME BETWEEN R WAVES:	400

TRIAL	BEATS	BAD BEATS	MEAN IBI	VAR	STD DEV	MEAN BPM
1	12	1	803.29	733.33	27.08	74.69
2	13	0	777.08	527.36	22.96	77.21
3	12	1	775.00	734.00	27.09	77.41
4	12	0	787.36	727.45	26.97	76.20
5	12	1	804.09	654.11	25.57	74.61
6	11	1	821.33	278.50	16.68	73.05
7	12	0	828.09	394.70	19.86	72.45
8	12	0	814.36	415.79	20.39	73.67
9	13	0	819.58	585.18	24.19	73.20
10	12	1	785.40	1997.38	44.69	76.39
GRAND STATISTICS			801.49	981.40	31.32	74.86



INTER-BEAT-INTERVALS



HE-86-9-117

EYEBLINK

EYEBLINK

1 NUMBER OF 10 SEC TRIALS (1-30)	3
2 EXTERNAL TRIGGER (Y/N)	N
3 SAVE RAW DATA? (Y/N)	Y

21 END
22 ABORT

EYEBLINK ANALYSIS

TRIAL NUMBER 3

NUMBER:	2	3
ONSET TIME:	153	941
AMPLITUDE:	2038	2026
DESCENT TIME:	5	11
CLOSING DUR.:	17	20
.50 WIN. DUR.:	14	18

CRITERION:

AMPLITUDE:	340
DURATION:	30
DROP:	150
WINDOW:	20

BLOCK RESULTS

TOTAL BLINKS: 3

MEAN CLOSING DURATION	18	MEAN BLINK INTERVAL (SEC)	8.075
		INTERVAL VARIANCE	0.07
MEAN .50 WIN. DURATION	14	INTERVAL STD. DEVIATION	0.27

NUMBER:	2	3
ONSET TIME (MSEC):	1530	9410
AMPLITUDE (MICRO VOLTS):	1	1
DESCENT TIME (MSEC):	50	110
CLOSING DUR. (MSEC):	170	200
.50 WIN. DUR. (MSEC)	140	180

HE-86-9-116

NWTB APPLICATIONS

- SIMULATOR STUDIES
  - — B52
  - — F16
  - — A10
- LABORATORY STUDIES
  - — CRITERION TASK SET
  - — DRUG EFFECTS
  - — PERFORMANCE TASKS
- FIELD STUDY
  - — A7

HE-86-9-98

NWTB AS A LABORATORY TOOL

- SEVERAL LAB TASKS USED TO PROVIDE STIMULI
- NWTB USED TO MEASURE CENTRAL AND PERIPHERAL RESPONSES
- CORRELATED SIMULTANEOUSLY RECORDED PHYSIOLOGICAL PERFORMANCE AND SUBJECTIVE DATA
- PHYSIOLOGICAL DATA TAPS SEVERAL ASPECTS OF "WORKLOAD"

NWTB APPLICATIONS

- LABORATORY
- VALIDATION STUDY
- AIR FORCE
- ARMY
- NAVY
- CDE

HE 85 5 34

W O R K L O A D

A psychophysiological approach

J. A. Stern, Ph.D.

BEHAVIOR RESEARCH LABORATORY

WASHINGTON UNIVERSITY

ST. LOUIS, MISSOURI

DEFINE WORK LOAD

IT CANNOT BE DEFINED SOLELY ON BASIS OF:

1. TAŞK DEMANDS - Input requirements
   - (Domain of engineer)
2. PROBLEM SOLVING STRATEGY - Decision making requirements
   - (Domain of psychologists)
3. RESPONSE COMPLEXITY - Output requirements
   - (Domain of human factors)

WORK LOAD IS:

The resultant of the interaction between:

1. TASK DEMANDS
2. PERFORMER ATTRIBUTES
3. RESPONSE REQUIREMENTS

WORK LOAD

AVERAGE - Integrated over time - MINUTES, HOURS

MOMENTARY - Integrated over time - SECOND(S)

OUR FOCUS - on MOMENTARY

WHY:

1. TASK DEMANDS - generally vary over time

2. PERFORMER ATTRIBUTES - generally vary over time

   MOTIVATION, ATTENTION, etc., FLUCTUATES

3. RESPONSE REQUIREMENTS - vary over time

How long is a MOMENT

Depends on question being asked

Work load associated with touch down and roll-out of aircraft

10 - 60 seconds

Work load concern with HUMAN ERROR

.1 - 30 seconds

HUMAN ERROR - leads to ACCIDENT - some of the time

CO-OCCURRENCE of

1. HIGH TASK DEMANDS

   Such as unusual environmental event

2. EQUIPMENT PROBLEM

   Such as blow-out of tire

3. HUMAN PROBLEM

   Attention - Distracted
   - Fatigued
   - Drugged
   - etc.

A MOMENT OF LOWERED ATTENTION coupled with MOMENT of HIGH TASK DEMANDS and/or EQUIPMENT PROBLEMS results in HUMAN ERROR which can result in ACCIDENT.

Our strategy for using physiological measures in work load assessment -

1. Laboratory studies to define physiological parameters that may reflect components of work load (attention, drowsiness, distracted, etc.)

   - RELATE physiological measures to PERFORMANCE measures

   - DEVELOP predictive equation allowing for use of physiological measure to predict alterations in PERFORMANCE

   - VALIDATE equation in a number of laboratory settings

2. Apply in SIMULATION or REAL WORLD

3. Develop hardware/software for real time implementation of measures.

What can physiology contribute to WORK LOAD assessment

1. It is resultant of interaction between

   A. TASK DEMANDS

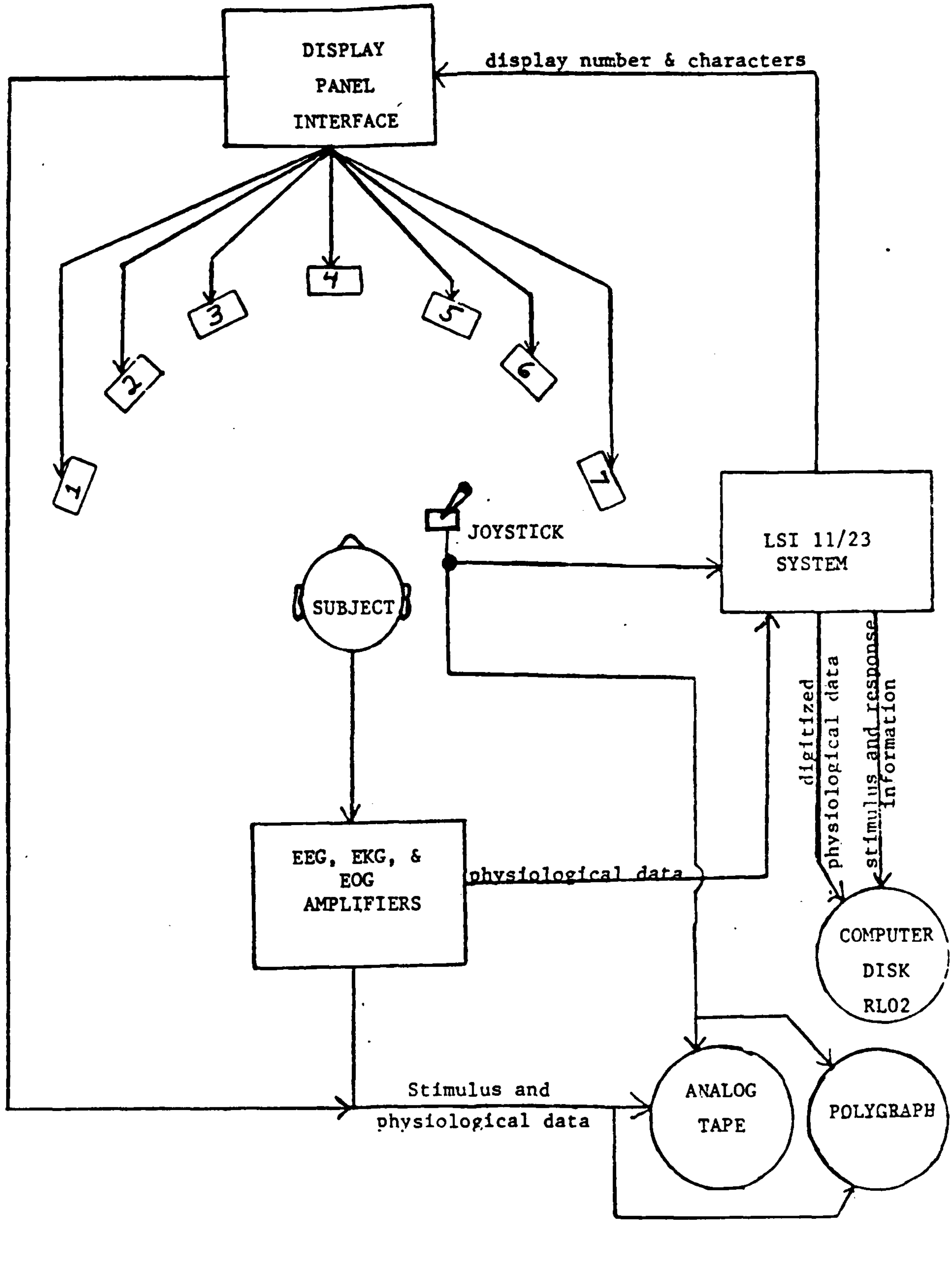
   B. PERFORMER ATTRIBUTES

   C. RESPONSE REQUIREMENTS - as recorded from the performer

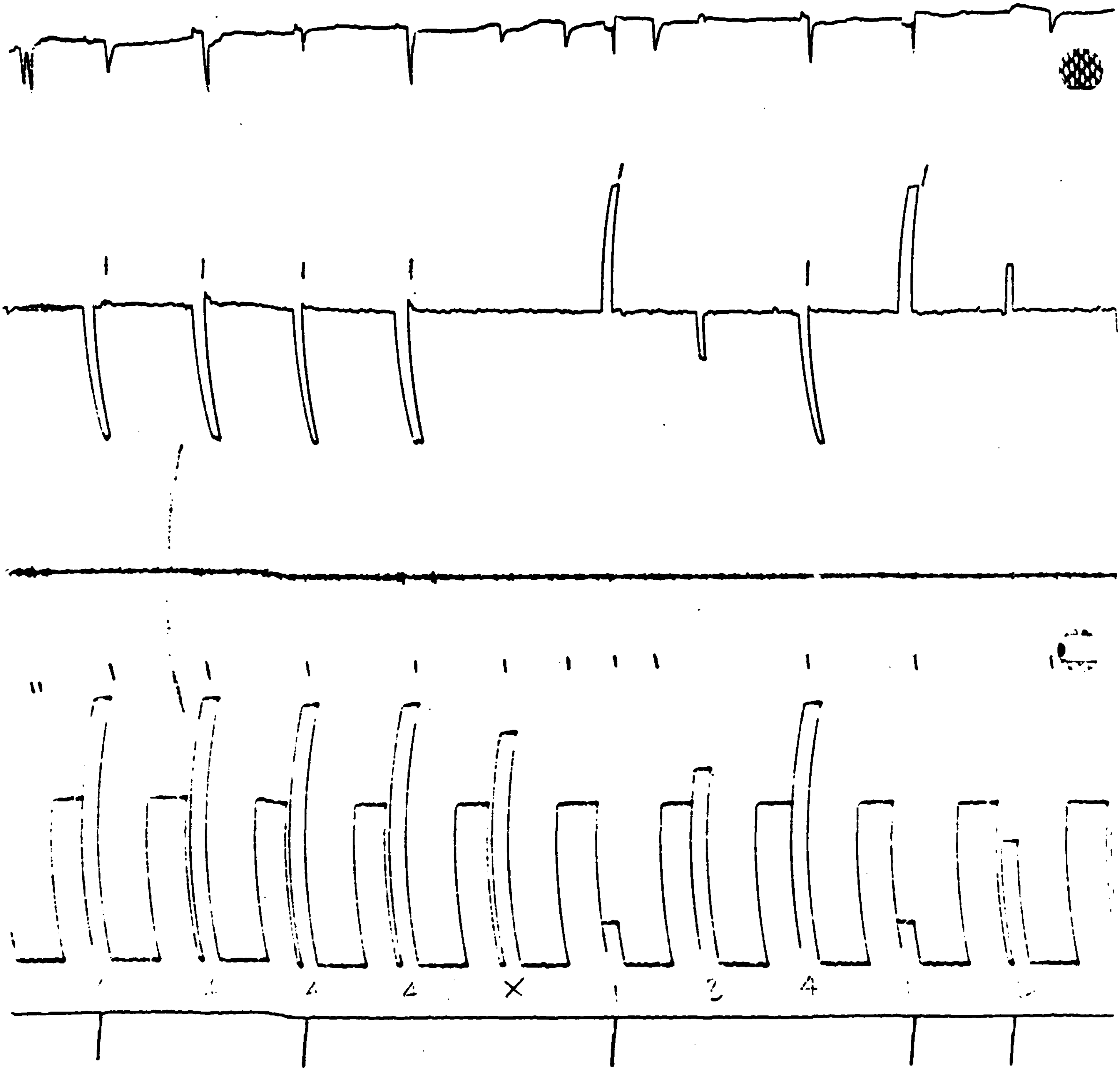
2. It can be sampled continuously

3. It can be sampled surreptitiously without interfering with task performance

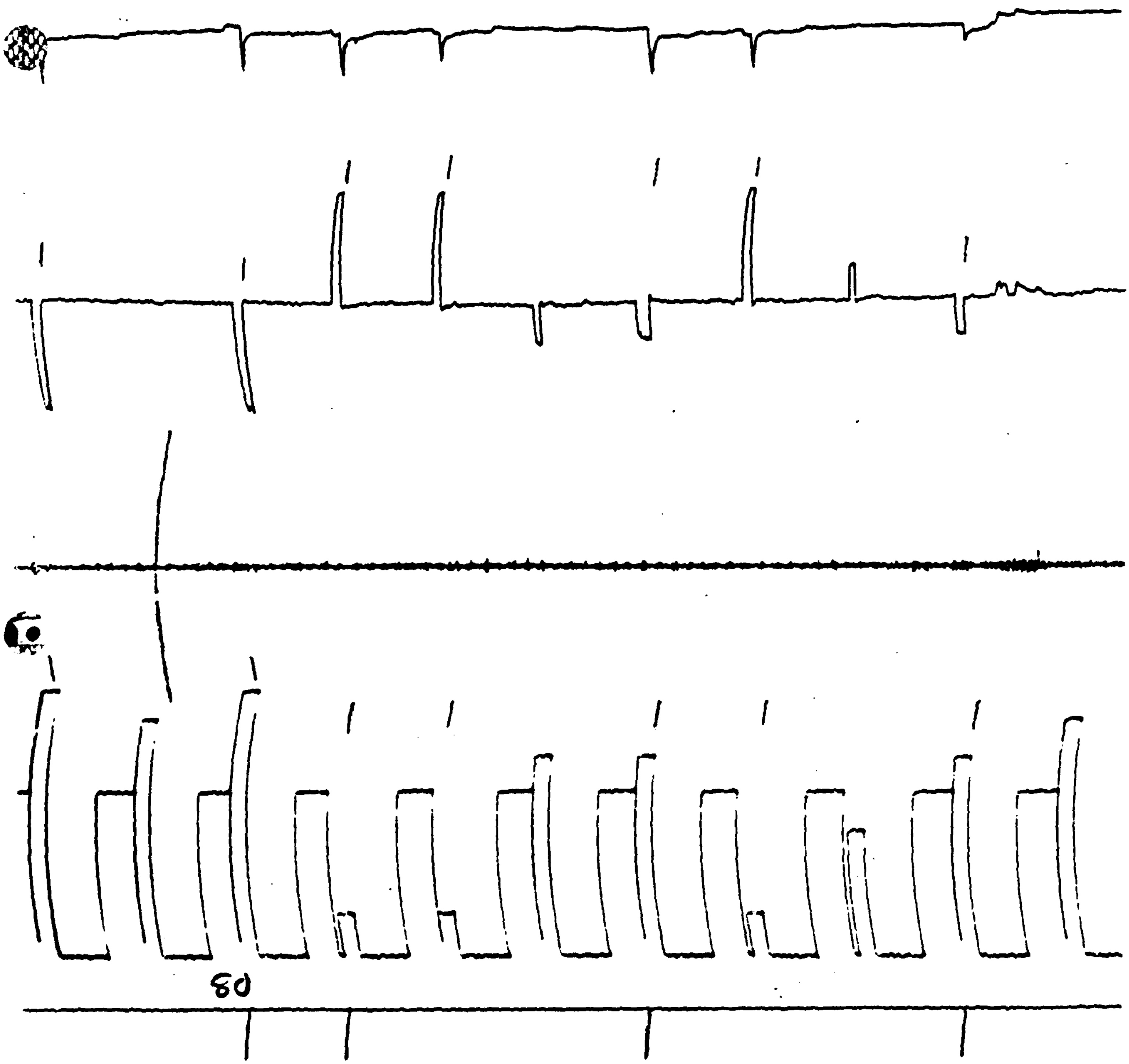
4. It is objective - not dependent on subject self-report.

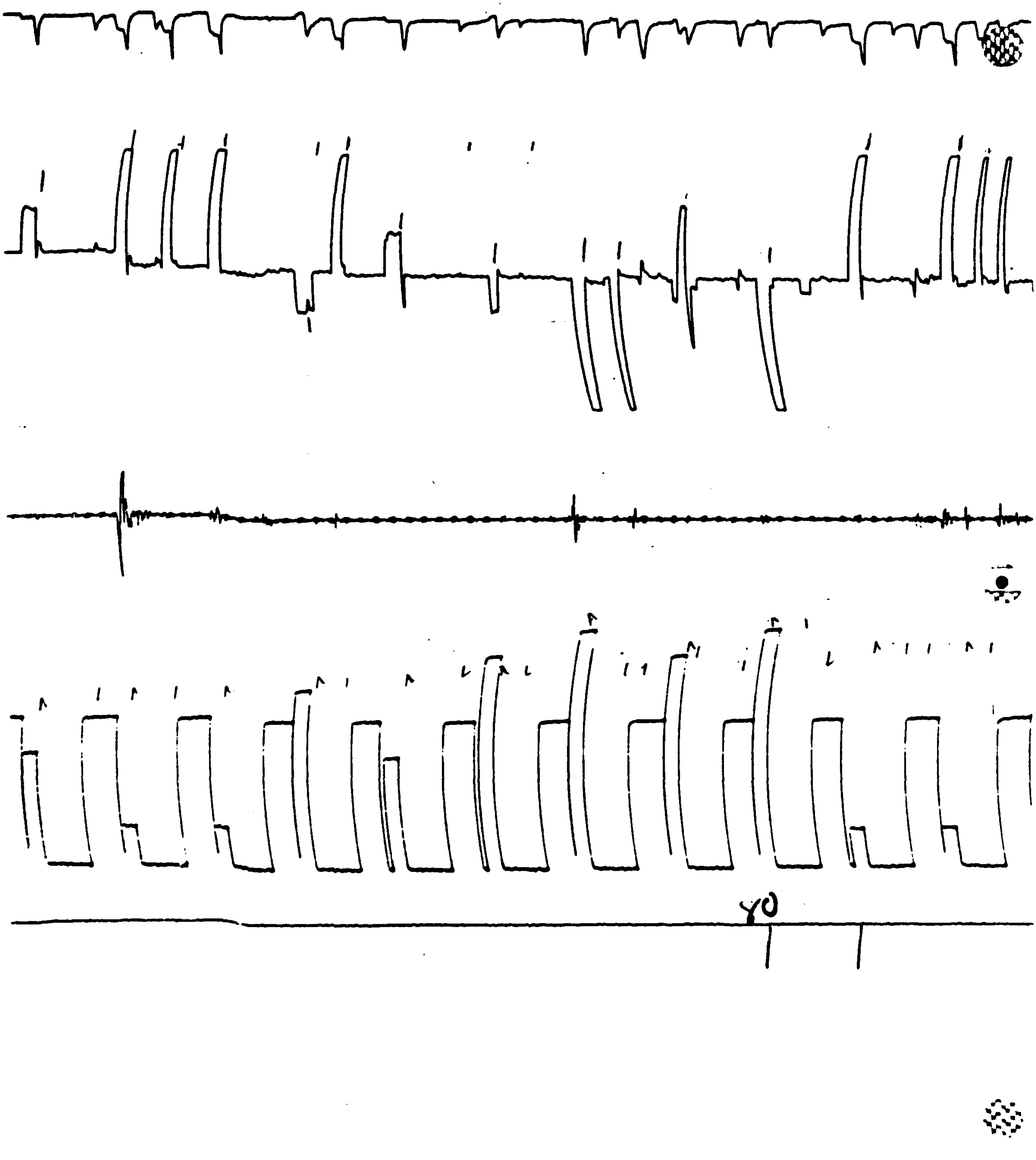

DISPLAY
PANEL
INTERFACE
display number & characters
1
2
3
4
5
6
7
JOYSTICK
LSI 11/23
SYSTEM
SUBJECT
digitized
physiological data
stimulus and response
information
EEG, EKG, &
EOG
AMPLIFIERS
physiological data
COMPUTER
DISK
RL02
Stimulus and
physiological data
ANALOG
TAPE
POLYGRAPH

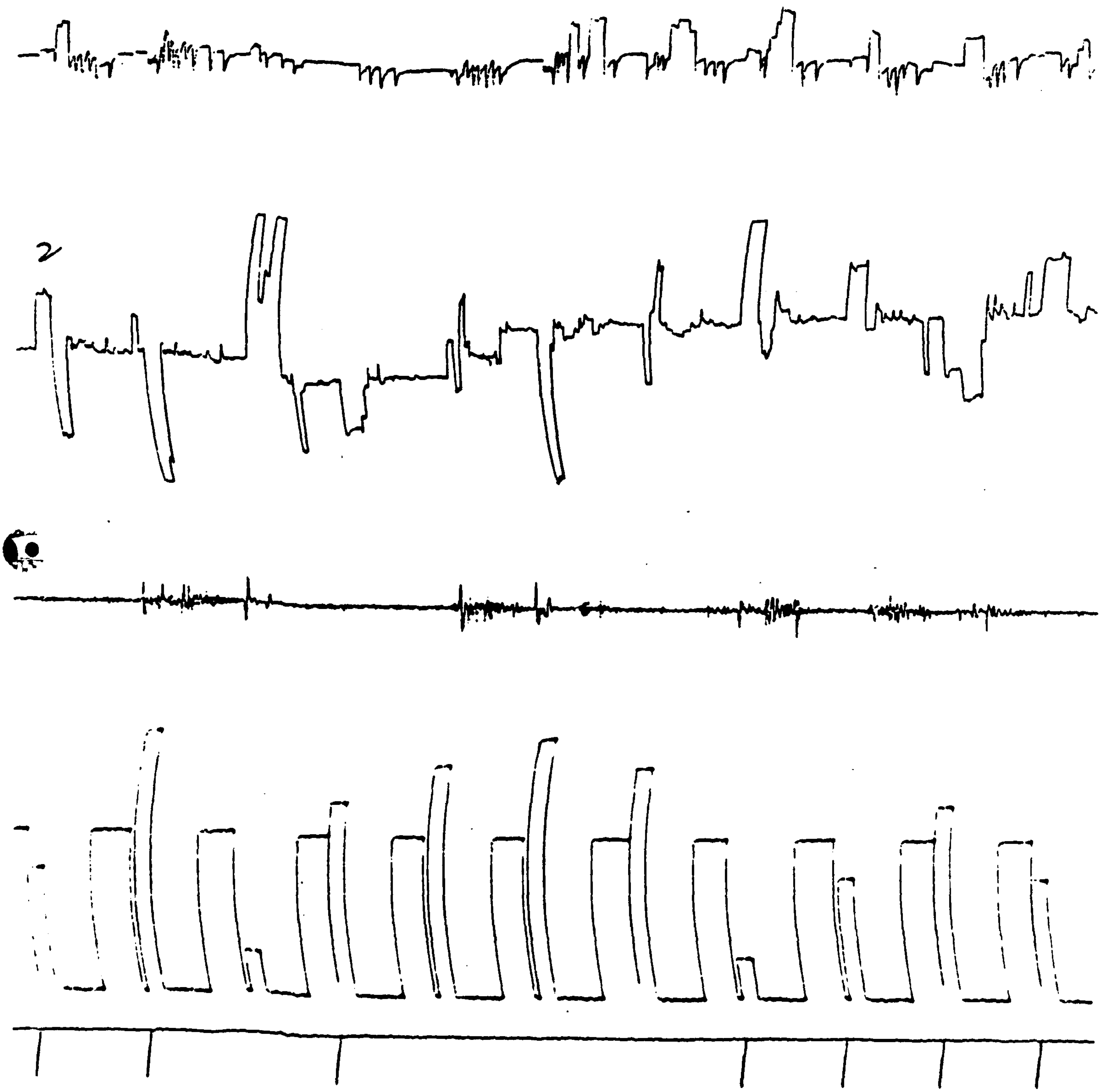
A

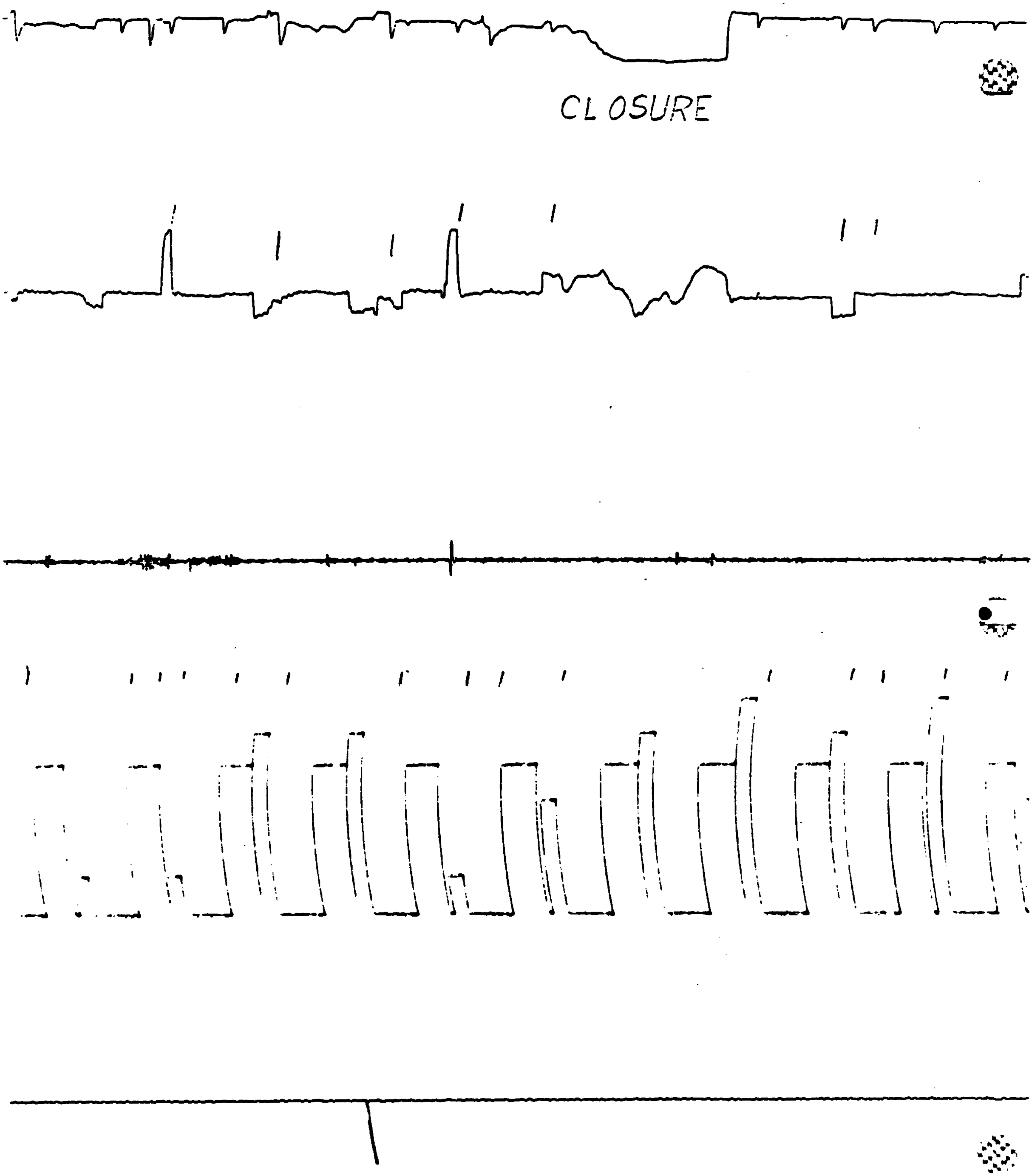


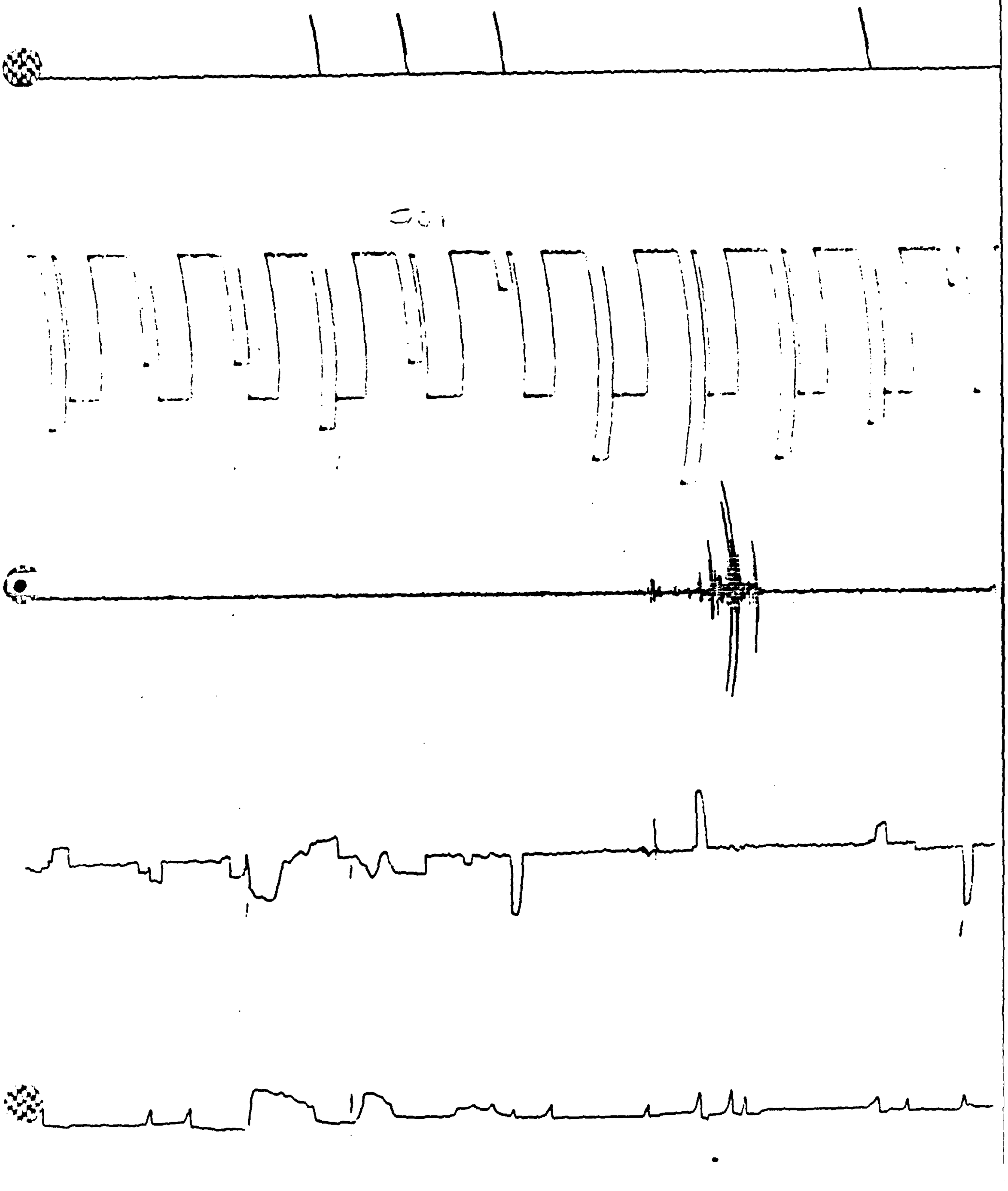
B

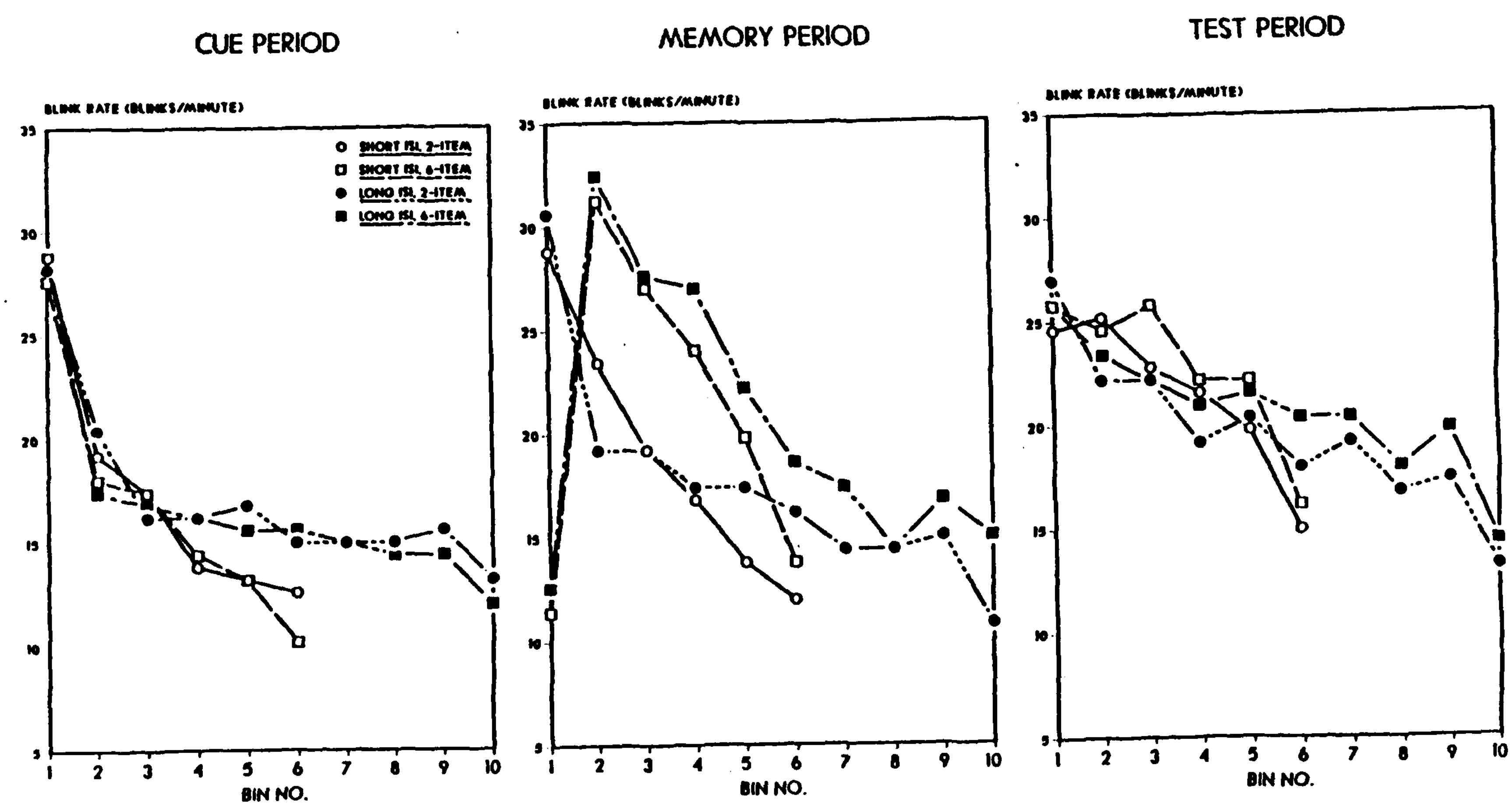


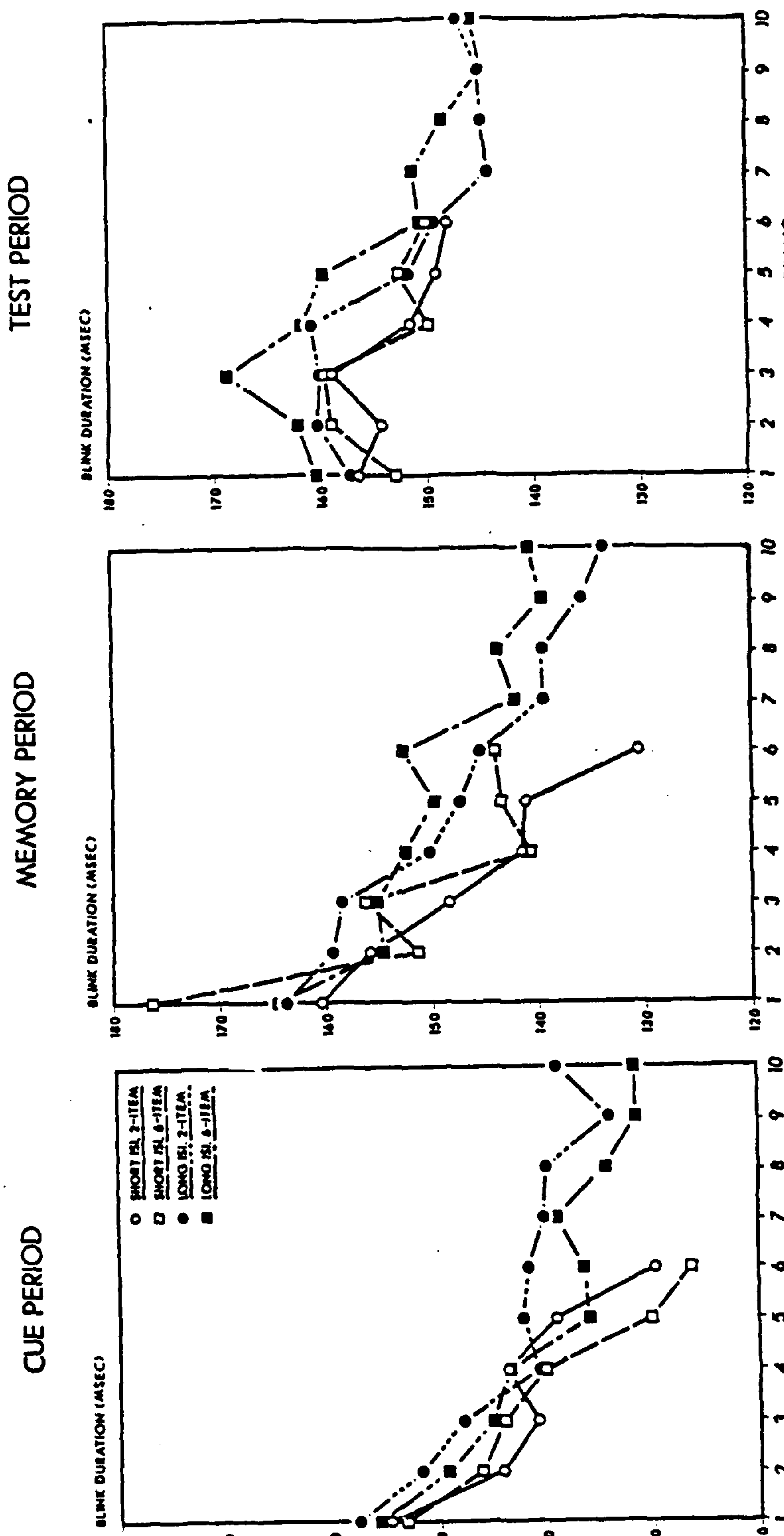

80

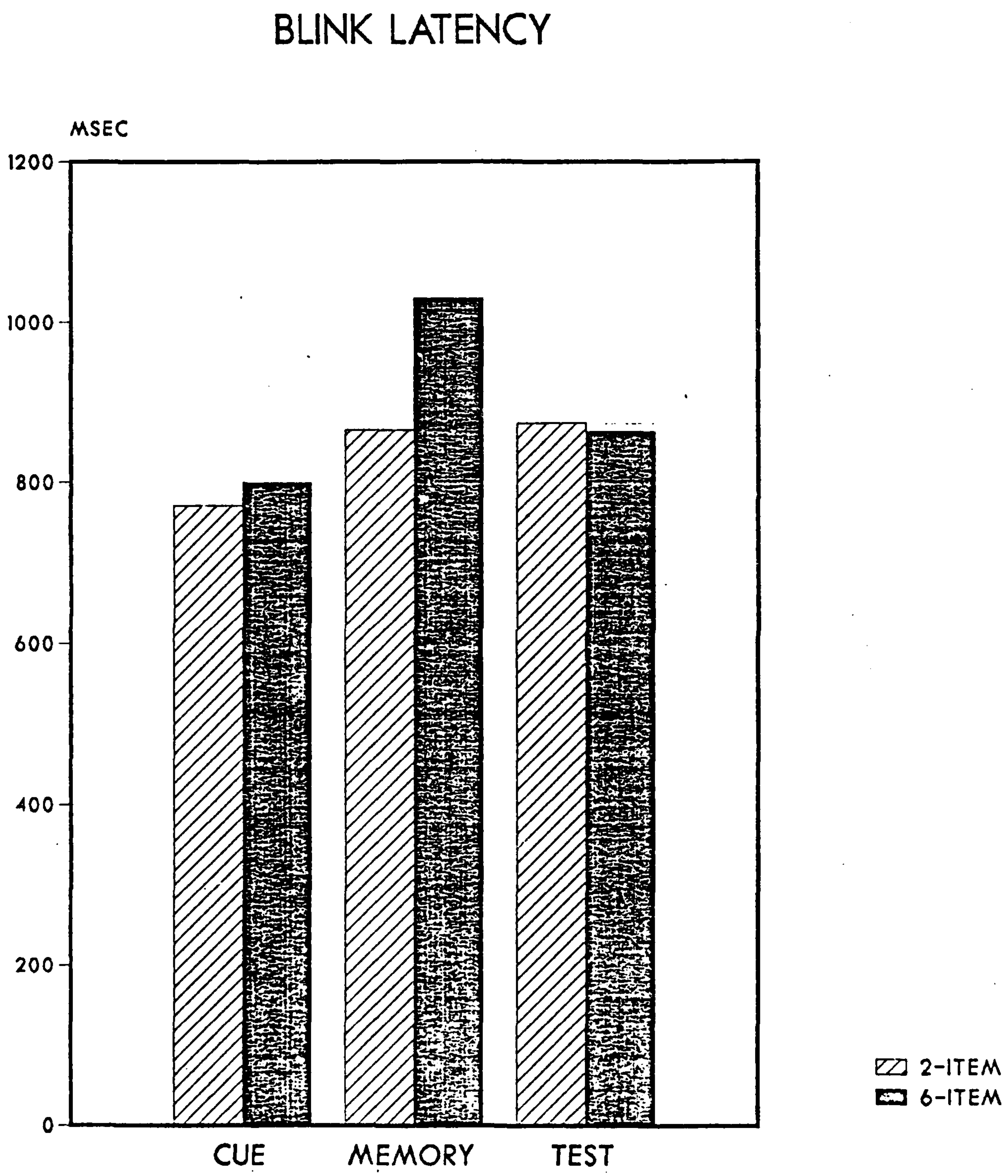

2


CLOSURE

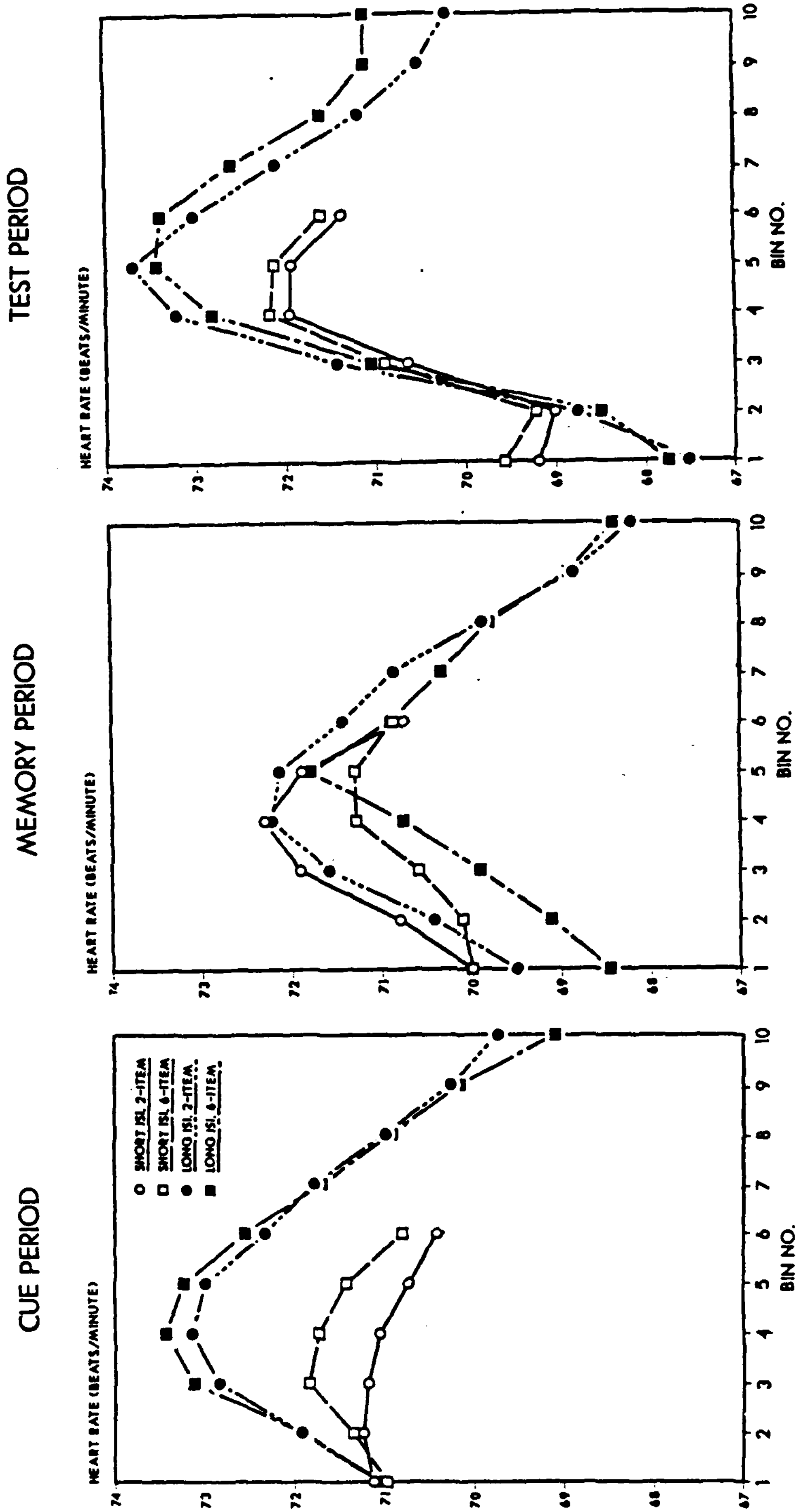



CUE PERIOD
MEMORY PERIOD
TEST PERIOD
BLINK RATE (BLINKS/MINUTE)
SHORT ISI, 2-ITEM
SHORT ISI, 6-ITEM
LONG ISI, 2-ITEM
LONG ISI, 6-ITEM
35
30
25
20
15
10
5
1
2
3
4
5
6
7
8
9
10
BIN NO.


CUE PERIOD
MEMORY PERIOD
TEST PERIOD
BLINK DURATION (MSEC)
BIN NO.
SHORT ISI, 2-ITEM
SHORT ISI, 6-ITEM
LONG ISI, 2-ITEM
LONG ISI, 6-ITEM
180
170
160
150
140
130
120
1
2
3
4
5
6
7
8
9
10


BLINK LATENCY
MSEC
1200
1000
800
600
400
200
0
CUE
MEMORY
TEST
2-ITEM
6-ITEM

313/314 Reverse Blank



SESSION 4. BEHAVIORAL TECHNIQUES

Presenters

Lt. Col. W. L. Derrick and Lt. Col. R. M. McCloy, USAF:
Behavioral Workload Assessment Techniques

BEHAVIORAL WORKLOAD ASSESSMENT TECHNIQUES

(Secondary Task Methodology)

William L. Derrick & Thomas M. McCloy
United States Air Force

Sources for this Presentation:

1) Handbook of Perception and Human Performance (K. Boff & Kaufman, Eds.), 1986.

a. Gopher, D. & Donchin, E. "Workload -- An Examination of the Concept", especially sections 1, 2.1, 2.5 - 2.8, and 3-3.8.

b. O'Donnell, R. & Eggemeier, T. "Workload Assessment Methodology", especially sections 1, 3, and 4.

2) Engineering Psychology and Human Performance, C. D. Wickens, 1984, especially chapters 1 and 8.

You should read these references. If you have little or no background in behavioral science, start with the O'Donnell and Eggemeier chapter. The Gopher and Donchin chapter is well written but it deals with some complex concepts. The Wickens text, which discusses the research and theories the other two readings draw upon, is graduate-level material.

You should also stay current in the literature. The Proceedings from the Annual Meeting and Human Factors have several articles on workload. But remember, almost anyone can publish anything in the Proceedings, so some questionable work may be found there. Also, Human Factors may have articles that are difficult to understand, deal with very theoretical (and esoteric) concepts that will not help a practitioner, and may contain work of a so-so quality even though the articles receive a peer review. Keep in mind that technical reports from laboratories and contractors may not receive much of a technical review either.

I. Why don't we have a "How-to" workload guide for secondary tasks?

* Wierwillie and Williges published guides in 1978

* The Boff and Kaufman volume is as close as we come in 1986.

* Reason: Several theoretical issues are still being resolved (and some may never be).

* Any guide (or research) that ignores these issues is snake oil.

II. Must start with a definition of workload. This will guide the selection of secondary tasks, their method of employment, and interpretation of results.

* Still won't find complete agreement on the definition, but we've made much progress.

* Can't define workload as just task difficulty. Any method of task description does not account for different people performing the task at different times in combination with other sets of activities. Any one task characteristic can produce a whole range of "workload" results.

* Can't just check system performance. Very different system designs can lead to equivalent criterion performance but intuitively we know their "workloads" are different.

* Can't just ask operators. Raters may be unreliable, have biases, recall difficulties, and poor (or no) insight into what produced the workload.

* Therefore, let's define workload as the interaction between the operator and the task. When the demands of the task exceed the operator's capability to deliver, we invoke the hypothetical construct of workload (or overload).

III. What do we ignore in this workload definition?

* motivation; assumed to be adequate.

* learning; the skills to perform the task are in the operator's repertoire.

* ability; if the operator does not possess the needed ability(ies), workload cannot explain the results.

IV. What operator capabilities must we consider?

* "Mental" workload suggests mental capabilities.

* Long and rich history of studies in experimental psychology suggest that we can conceptualize an information processing system between task stimulus and operator response.

* This system has an architecture (e.g., perception, memory, response selection) and an energetical component. Terms such as attention, mental energy, processing resources, processing capacity refer to the latter.

* "mental workload may be viewed as the difference between capacities of the information-processing system that are required for task performance to satisfy performance expectations and the capacity that is available at any given time." (Gopher & Donchin) This is a closed-loop definition.

* Task difficulty may now be defined by amount of resources demanded. A "different" task demands most of the resources available; a "very difficult" or "overly difficult" task demands more resources than are available at that instant in time.

V. What are some of the characteristics of processing resources?

* the fuel that makes the information-processing system operate.

* limited in quantity at any point in time.

* probably not expanded by arousal.

* can be allocated to different parts of the architecture.

* not the same thing as the laymen's concept of attention; not open to introspection (thus a problem if subjective measures are used).

* experimental data suggests more than one type or pool of resources (this complicates workload assessment considerably).

* not possible to observe so must make inferences based upon task performance; however, lots of possible reasons for any given level of task performance.

* research suggests that certain highly skilled tasks with specific input and response characteristics may demand almost no resources at all even though the task was resource demanding at one time (an "automatic" process).

VI. So what does all this have to do with the use of behavioral techniques in the assessment of workload?

* We desire to know the composite resource demands of the task of interest. We will call this workload and note that it is a function of an interaction between the operator and the task.

* Resource demands can be inferred only from the results of task performance.

* The task of interest most likely has multiple resource demands. We must uncover the resource load profile. Wickens states that a load profile will consist of three dimensions:

  1) stages of processing: perception & central processing vs. responding.

  2) codes of central processing: verbal vs. spatial.

  3) modalities of input and response: auditory vs. visual input; manual vs. vocal response.

* We look at the primary task and make some guesses as to what resources are demanded for acceptable task performance. We then choose a battery of secondary tasks that are known to demand specific types of resources. If we are lucky, we know the ordinal amount of resources demanded by these secondary tasks (little, some, a lot). The primary tasks must interfere on some of the resource dimensions.

* When tasks "interfere" they demand the same resources for performance. If the joint resource demand is great enough, some degredation in primary or secondary task performance (or both) should occur. This will permit you to estimate what type and roughly what level of resources are demanded by the primary task.

* In general, this approach is diagnostic. It tells you what types and what levels of resources are demanded by the primary task when performed to criterion. Thus, you obtain a resource load profile (workload profile) on three dimensions.

VII. More specifically, what steps are employed?

1) The task of interest is isolated and analyzed for potential resource demands.

2) Operator motivation is established at the appropriate level for the user population.

3) The task is learned by the operators until performance reaches a plateau or some level of stability. Good luck.

4) The battery of secondary tasks is selected. Hopefully none are locally developed. These tasks should demand some of the same resources as the primary task and those resource demands should have been previously validated.

5) Each secondary task is practiced until a stable level of performance is reached.

6) The primary and secondary tasks are performed jointly. Each secondary task should have some aspect that interferes with the primary task but does not intrude.

7) Operators are given the instruction that the primary task should take priority; i.e., its performance in the dual task mode should equal performance in the single task situations (but see below).

8) Single-to-dual task performance changes for each secondary task are examined.

9) A pattern of performance decrements and no decrements should emerge. This pattern can be translated into a load profile. For example, you might conclude that your primary task has the following load profile (e.g., workload):

a. Some demand for perception and central processing resources and a large demand for response-execution resources.

b. Some demand for both verbal and spatial resources.

c. No demand for auditory resources but a very heavy demand for visual resources. Somes demand for manual but a very small demand for vocal response resources.

VIII. What might produce nearly perfect dual task performance?

* The composite set of resource demands in the primary task has nothing in common with the resource demands in the secondary task. Unlikely if you properly analyzed your primary task and chose a validated secondary task.

* The resource demands of the primary or secondary tasks (or both) were "data limited" (translation: very small separate demands so that the combined demand could easily be accomodated by the operator). A real possibility. If your secondary task is data limited, you do not have a diagnostic instrument.

IX. Where do I get my hands on these secondary tasks?

* If you have the theoretical background, the laboratory and equipment, a software person, and the time, you can create them yourself. Read about the tasks used in the resource model literature and build them for your use. Much can go wrong here. We don't recommend this approach for the practitioner with a workload answer to report in a relatively short period of time.

* A criterion task battery is being developed. Some papers will be reported at this conference. You may be able to get enough detail from the authors to create these tasks for your own use. They may be available for your use if you provide the data to the researchers.

* Consultants are always out there but be careful. You don't have to understand anything on the proceeding pages to call yourself a consultant.

* Do not use a set of tasks you happen to have leftover from previous work just because they are available. They must be validated as to the types and ordinal levels of resources demanded.

X. How should dual-task performance decrements be measured?

* The answer isn't an easy one to give. Debate exists as to how to scale the changes in performance from single to dual task and how many data points are needed for each primary-secondary task pair.

* One procedure, developed by Wickens is based upon the normalized score concept (score-mean/standard deviation). For any task, the mean = the operator's single task performance, the score = the operator's performance of that task when in the dual task mode, and the standard deviation = the single task performance standard deviation. This procedure produces a decrement score that can be compared to a single task score. Decrement scores for all tasks can be plotted for visual comparison in what has been called a POC (Performance Operating Characteristic) space. (However, this is not a true POC - see below). Statistical tests (very complicated ones) can compare all dual task points to determine if the decrements are significant (See Wickens et.al., 1980, in Human Factors and Derrick, 1981, in Proceedings for examples).

* Another approach is simply to scale performance on the metric appropriate for each task (e.g., task 1 - reaction time, task 2 - RMS error). Changes within each task can be evaluated statistically, but evaluation of joint task performance is not clear since you have the old apples vs. oranges problem (See Gopher et.al., 1982 in Journal of Experimental Psychology: Human Perception and Performance, and our article in the 1984 Proceedings).

XI. What do you mean by number of data points in X above?

* A very good case can be made against the classic dual task methodology considered thus far. Recall that operators are asked to guard or protect primary task performance in the dual task situation. The purpose here is to index the types and levels of resources demanded for normal primary task performance with performance changes (or lack thereof) in the battery of secondary tasks. This is probably a naive approach. First, as the dual task is practiced, strategies may be developed to integrate the separate task components, so the joint resource demand is less than the sum of the separate resource demands. Second, a type of dual task interference may arise that produces a drop in performance that has nothing to do with the true resource - workload question (e.g., increased visual scanning). Wickens originally termed this structural interference. Third, operators may not follow the instructions to guard primary task performance; it may drop or even increase, and you are stuck with a secondary task score that in and of itself gives you erroneous information.

* The answer to these problems is the true POC, not the single point POCs you will find in Wickens et.al., 1980 and Derrick, 1981, Gopher et.al., 1982, first reported this although ours (1984) utilizes a better methodology. Here operators are told to allocate X percent of their "attention", a term they can understand, to one task and 1-X percent to the other. At least three different allocation percentages per dual task are used. The resultant data produces some type of curve that depicts the whole range of resource and performance tradeoffs for a primary-secondary task pair. Thus you get the resource "cost" for any given level of performance, and that relationship may be anything but linear.

* POCs are all well and good theoretically, but they are very difficult to obtain. The allocation policy instructions require the clever use of on-line performance feedback if they are to be followed. To get the needed data from an operator requires 3 to 5 times the laboratory time for each primary - secondary task pair of interest.

XII. Will operators accept the use of secondary tasks to assess workload?

* The evidence thus far isn't good. First, how can you describe workload in terms of processing resources to an operator which must be done to justify the method? It sounds like magic or nonsense. Second, the types of tasks themselves have no "face validity" for what the operator considers to be workload; therefore, expect resistance. Third, to implement the required task battery yourself takes a lot of time, equipment, and special skills. Fourth, to conduct the secondary task workload study requires a great deal of operator time and skill on your part. You should never attempt secondary task work unless you have been an apprentice. Fifth, proper analysis of the data is by no means straightforward. Any given result may have many possible and interconnected causes.

* Embedded secondary tasks have been proposed as an alternative. The problem here is that what can be embedded may well be unique to each workload situation. Thus, the embedded task may not be validated as to its resource demands, no reliability data may exist, it would be difficult to use an allocation procedure to get a POC with an embedded task, etc.

XIII. So what's the bottom line on all of this?

* The good news is that many of the "leading lights" in the workload area are coming to the position that workload must be defined as resources consumed during task performance, they agree with Wickens on the nature of those resources, they have identified the proper methods of secondary task workload assessment, and they are developing a validated battery of secondary tasks to measure resources.

* The bad news is that the theoretical issues here are complex and almost require graduate training in cognitive psychology, the process itself is time-consuming, frought with peril, and requires skills just beneath that of brain surgery, operators don't like it, and results are difficult to interpret and even more difficult to explain to a project engineer.

* Should you give secondary tasks a chance to answer your next workload question? Only if you are willing to read about the issues, practice the method where it doesn't count, and then invest the time and money to do it right. Otherwise, hire a consultant with a track record of correct secondary task workload assessment. If you choose to disregard this approach all together and use a different one, realize that other approach is unlikely to give you the complete workload picture.

SESSION 5: OTHER TOOLS

Presenter

R. A. North: A Workload Index for Iterative Crewstation Evaluation (W/INDEX)

A WORKLOAD INDEX FOR ITERATIVE CREWSTATION EVALUATION

Robert A. North

Honeywell Systems and Research Center, Minneapolis, MN

ABSTRACT

A crewstation design tool is reviewed which allows the human factors engineer to assess the attentional demands that will be imposed on the human operator given the tasks, times of performance, time-sharing demands, individual task difficulties, and human interfaces to be used. Attention and performance theories used in the workload model are discussed. A crewstation design problem example is used to illustrate the utility of the tool in pointing to automation needs and potential reallocation of tasks to display/control surfaces.

INTRODUCTION

Often, the crewstation design team is faced with decisions about assignment of crew tasks to displays and controls early in the design process without a method of evaluating the overall design's impact on the attentional resources, and subsequent workload demands on the human opertors. Mission and task analysis, knowledge of human performance, and initial selection of generic display/control areas are enough to begin gaining valuable insight into eventual design problems, workload peaks, and automation needs. A procedure for combining these data into an interactive, iterative design tool has been developed to provide a workload profile across time. The metric, known as the Workload Index (W/INDEX), has been used on several recent advanced design programs, as well as for evaluation of existing crewstations. In this paper, the metric's features will be reviewed and illustrated with a design example.

W/INDEX METHODOLOGY

W/INDEX is a combination of mission, task, and timeline analysis, and theories of attention and human performance used to predict attentional aspects of workload in the crewstation. Previous models such as Timeline Analysis discussed by Parks (1979) include the timeline modelling and workload estimation, but do not include a way to estimate the differential effects of time-sharing load caused by overlapping resource demands. This latter issue is the primary difference between W/INDEX and other techniques. W/INDEX computes a level of workload demand for each second of a mission by assessing individual task difficulties and the special problems of time-sharing between activities. This computer-based technique operates on the following data:

1. Crewstation Interface Channels which represent input/output of information between the vehicle systems and the operator. In initial crewstation design phase, this list may represent generic categories (e.g., visual, manual, verbal, auditory) and may be expanded to include specific display/control surfaces as the design matures.

2. Human Activity List representing the present or anticipated tasks to be performed in the course of the mission/flight.

3. Attention Involvement Levels which estimate the relative attention demand of each interface that is required to perform a given crew activity. This relative scaling is obtained by operator opinion, or any existing simulation performance data.

4. Interface Conflict Matrix specifying the time-sharing penalties associated with simultaneous demands on attention between two crewstation interface channels (e.g., displays ro controls). This includes problems arising from suboptimal placement of displays/controls, and human capabilities and limitations in time sharing between certain channel combinations. The interface conflict feature reflects models of attention and time sharing known as "resource limitations" being researched and validated by Dr. Christopher Wickens and his staff of the University of Illinois. These attention models reflect the degradation in human performance attributed to sharing similar resources and improvement when dissimilar resources are shared (Wickens, 1984).

5. Operator Activity Timelines, which specify the start and stop times of each activity during the mission or flight. Times must be derived from existing procedural doctrine, or if procedures do not yet exist, from estimations using similar types of tasks. Part task simulation testing can be used in deriving these task times.

The computation of the workload demand index operates on the above data sources to produce a second-by-second profile of demand. Peak demands may then be examined in more detail to determine possible workload alleviation methods (e.g., information fusion, use of voice activation to offload manual channel, etc.). This method provides an iterative tool for evaluating the workload impact of cockpit design modifications.

A generic example based on an airborne crewstation scenario illustrates the technique. Table 1 lists the human channels or interfaces which represent generic categories that can be expanded in later design phases. Table 2 shows the crew activity list, and estimated attentional involvement per activity and display control interface. A fixed-wing, single-seat air-to-ground target engagement (anti-armor) with a "low-and-fast" terrain following profile was assumed for the example. Three levels of automation are illustrated: (1) manual flight and targeting (no automation), (2) automated target search/identification, and (3) automated flight and target search/identification.

The numbers in the cells of the matrix in Table 2 represent a difficulty index for human performance indicating an estimate of the attentional involvement required by that task/channel combination, on a 1-5 scale. We have experimented with various methods of obtaining this

TABLE 1. NO AUTOMATION - BASELINE COCKPIT

TASK		WIN	HMD	CRT	AUD	STK	KEY	DIS	SPE	COG
1	200' AGL manl	5	2	0	0	4	0	0	0	0
2	200' AGL auto	3	1	0	0	0	0	0	0	0
3	Rcv tgt data	0	0	0	2	0	0	0	0	0
4	Ack tgt data	0	0	0	0	0	0	0	1	0
5	Enter tgt data	0	0	2	0	2	0	0	0	0
6	Swtch freqs	0	2	0	0	2	0	0	0	0
7	Engage atoplt	0	1	0	0	0	0	0	0	0
8	Survey enemy	3	0	0	0	0	0	0	0	0
9	Priortze tgts	0	0	0	0	0	0	0	0	3
10	Select weapon	0	0	0	0	0	0	1	0	0
11	Arm weapon	0	0	0	0	0	0	1	0	0
12	Track target	0	3	0	0	0	0	0	0	0
13	Launch weapon	0	0	0	0	0	0	1	0	0
14	Verify track	0	2	0	0	0	0	0	0	0
15	Assess damage	2	0	0	0	0	0	0	0	0

TABLE 2. AUTOMATED TARGETING - SPEECH TECHNOLOGY

TASK		WIN	HMD	CRT	AUD	STK	KEY	DIS	SPE	COG
1	200' AGL manl	5	2	0	0	4	0	0	0	0
2	200' AGL auto	3	1	0	0	0	0	0	0	0
3	Rcv tgt data	0	0	0	2	0	0	0	0	0
4	Ack tgt data	0	0	0	0	0	0	0	1	0
5	Enter tgt data	0	2	0	0	0	0	0	2	0
6	Swtch freqs	0	2	0	0	0	0	0	2	0
7	Engage atoplt	0	0	0	0	0	0	1	0	0
8	Init ATR	0	1	0	0	0	0	0	1	0
9	Priortze tgts	0	0	0	0	0	0	0	0	0
10	Select weapon	0	0	0	0	0	0	0	1	0
11	Arm weapon	0	0	0	0	0	0	1	0	0
12	Verify ATR tk	0	0	2	0	0	0	0	0	0
13	Launch weapon	0	0	0	0	0	0	1	0	0
14	Vfy weapon lc	0	2	0	0	0	0	0	0	0
15	Assess damage	2	0	0	0	0	0	0	0	0

number including pilot questionnaires with various anchor point descriptors (e.g., "simple vs. complex, "easy vs. difficult"). Other techniques such as the Subjective Workload Assessment technique might be used to derive this number. Aldrich and McCracken (1984) propose a technique which relates an index of difficulty to specific task descriptors (e.g., for cognition: 1=automatic (stimulus-response); 2=sign, signal, recognition; 3=selection from alternatives; 4=encoding/decoding, recall; 5=planning (projecting action sequence); 6=evaluation in decision making; 7=estimation, calculation, or conversion.

Table 4 illustrates the conflict matrix of time-sharing difficulty for pairwise combinations of the channels or interfaces. The scaling of the conflict weights is 0.0 to 1.0 which corresponds to parallel vs. single-channel performance capability as demonstrated in experiments by Wickens, Mountford, and Schreiner (1981) and others (Navon and Gopher, 1979; Kantowitz and Knight, 1976; and North, 1977), and represented by Wickens' Performance Operating Characteristic (POC). (The POC shows the results of performance when two tasks are time-shared in terms of their original single-task performance.)

W/INDEX APPLICATION

W/INDEX is most effective in several areas of the design process. It can be used in the initial stages of design to help spot problems with operator workload and find task elements that could benefit from automation. As the design matures, certain tradeoffs can be performed with the metric to check the effects of inserting different operator interface technologies such as speech command, helmet sighting/displays, multi-function keyboards, etc. In the following examples, these uses of the metric are illustrated.

Tables 1, 2 and 3 represent tasks and attention demand loadings for three fighter cockpit design concepts. Table 1 illustrates a non-automated, conventional type of design defined as a "baseline." Table 2 illustrates insertion of speech command technology for several tasks that were performed manually in the baseline (see SPE column), and automatic target recognition (ATR). Table 3 illustrates automation of aircraft control in the 200 ft. above ground level flight ingress phase and weapon delivery against armored targets.

The plot in Figure 1 illustrates the three profiles that resulted from applying the W/INDEX metric to the three cockpit concepts and the timeline data. The attention savings brought about by inserting speech command and ATR are evident in the dashed line profile in two areas. The early area between seconds 10 and 20 are caused by shifting the target data entry task to speech input. The later savings in attention (seconds 33 to 45) are due to automated target search and prioritization. A slight increase in attention is registered for Cockpit #2 in seconds 45 through 51 due to a shift of demand to the in-cockpit CRT for verifying the automated tracking process. If this were shifted back to the helmet display concept, this difference would disappear. The third cockpit adding automation of flight, registers a profile about 10 points lower than the baseline throughout the segment. Note that the automation of flight does not completely attenuate workload: Visual attention is still paid to the aircraft's altitude and outside world relationships.

The overall integrated attention scores for the three cockpits were 20.1 for the baseline, 17.9 for the cockpit #2 concept, and 8.7 for the fully automated version (cockpit #3). Thus we can estimate a "figure of merit" of pilot attention demand for the three cockpits using the W/INDEX metric.

TABLE 3. AUTOMATED TF/TA AND TARGETING

TASK		WIN	HMD	CRT	AUD	STK	KEY	DIS	SPE	COG
1	200' AGL manl	5	2	0	0	4	0	0	0	0
2	200' AGL auto	3	1	0	0	0	0	0	0	0
3	Rcv tgt dta	0	0	0	2	0	0	0	0	0
4	Ack tgt data	0	0	0	0	0	0	0	1	0
5	Enter tgt data	0	2	0	0	0	0	0	2	0
6	Swtch freqs	0	2	0	0	0	0	0	2	0
7	Engage atoplt	0	0	0	0	0	0	1	0	0
8	Init ATR	0	1	0	0	0	0	0	1	0
9	Priortze tgts	0	0	0	0	0	0	0	0	0
10	Select weapon	0	0	0	0	0	0	0	1	0
11	Arm weapon	0	0	0	0	0	0	1	0	0
12	Verify ATR tk	0	0	2	0	0	0	0	0	0
13	Launch weapon	0	0	0	0	0	0	1	0	0
14	Vfy weapon lc	0	2	0	0	0	0	0	0	0
15	Assess damaage	2	0	0	0	0	0	0	0	0

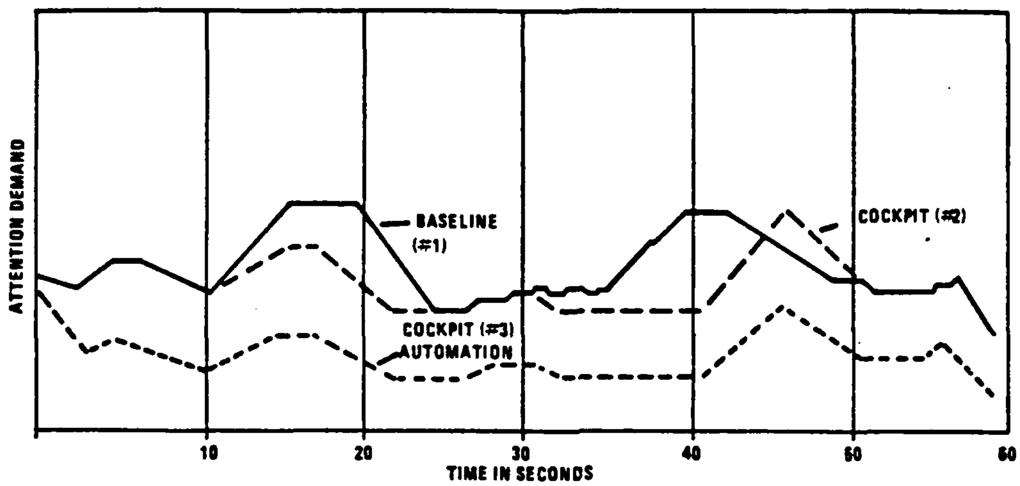


Figure 1. Comparison of Attention Levels for Three Cockpit Concepts

SUMMARY AND CONCLUSIONS

W/INDEX is a flexible tool for gaining quick insight into the potential attentional demands that different cockpit concepts will impose on an operator/pilot. The metric has three major uses:

1. Initial design stages to identify automation needs.

2. Intermediate design stages to estimate the effect of using different display and control concepts.

3. Applied to existing cockpits to look for most severe problem areas for redesign.

The metric has the following advantages over other methods of estimating workloads:

1. Rapid reconfigurations of cockpit designs may be evaluated by simple additions or editing of the database.

2. Metric considers the multiple-task management situation as a major component of workload, and uses human performance and attention models to estimate time-sharing difficulty.

3. Includes a cognitive element in estimating workload demand, taking into account decision making, memory recall, or other higher-order processes that the operator must call upon during performance.

The disadvantages of the metric are its lack of sensitivity to transient effects such as fatigue and stress, and its rigidity in modeling the time sequence of events in the mission scenario. This latter problem may be addressed through incorporating a method of "windowing" the period of time that each event may occur, for those tasks that are likely to shift their start and stop times in the mission. A "mean" and "standard error" concept could be applied to the data to introduce a stochastic component into the model and allow events to "occur" with an element of randomness.

A validity check on the predictive power of W/INDEX is the most critical need for further metric development. We are planning initial validations against simulation performance in armored and rotorwing crewstations during 1985.

REFERENCES

Gopher, D., & Navon, D. (1980). How is performance limited: Testing the notion of central capacity. Acta Psychologica, 46, 161-180.

Kantowitz, B. H., & Knight, J. L. (1976). Testing tapping time-sharing I: Auditory secondary task. Acta Psychologica, 40, 343-362.

North, R. A. (1977, October). Task functional demands as factors in dual-task performance. Paper presented at the 25th annual meeting of the Human Factors Society, San Francisco.

Parks, D. (1979). Current workload methods and emerging challenges. In N. Moray (Ed.), Mental workload: Its theory and measurement. New York: Plenum Press.

Wickens, C. D., Mountford, S. J., & Schreiner, W. (1981). Multiple resources, task hemispheric integrity, and individual diffcerences in time-sharing. Human Factors, 23, 211-229.

McCraken, J. H., & Aldrich, T. B. (1981) Analyses of selected LHX mission functions: Implications for operator workload and system automation goals. Anacapa Sciences, Inc.

TABLE 4. CONFLICT MATRIX OF TIMESHARING DIFFICULTY

TASK	WIN	HMD	CRT	AUD	STK	KEY	DIS	SPE	COG
Window (WIN)	.7								
Helmet Display (HMD)	.3	.3							
Cockpit CRT (CRT)	.9	.3	.5						
Auditory Input (AUD)	.3	.3	.3	.9					
Stick Control (STK)	.1	.1	.1	.1	.9				
Keyboard (KEY)	.1	.1	.1	.1	.7	.9			
Discrete Switch (DIS)	.1	.1	.1	.1	.5	.7	.7		
Speech (SPE)	.1	.1	.1	.3	.3	.5	.1	.9	
Cognitive (COG)	.5	.5	.5	.7	.3	.5	.1	.8	.9

CONFLICT MATRIX BETWEEN CHANNELS

Honeywell

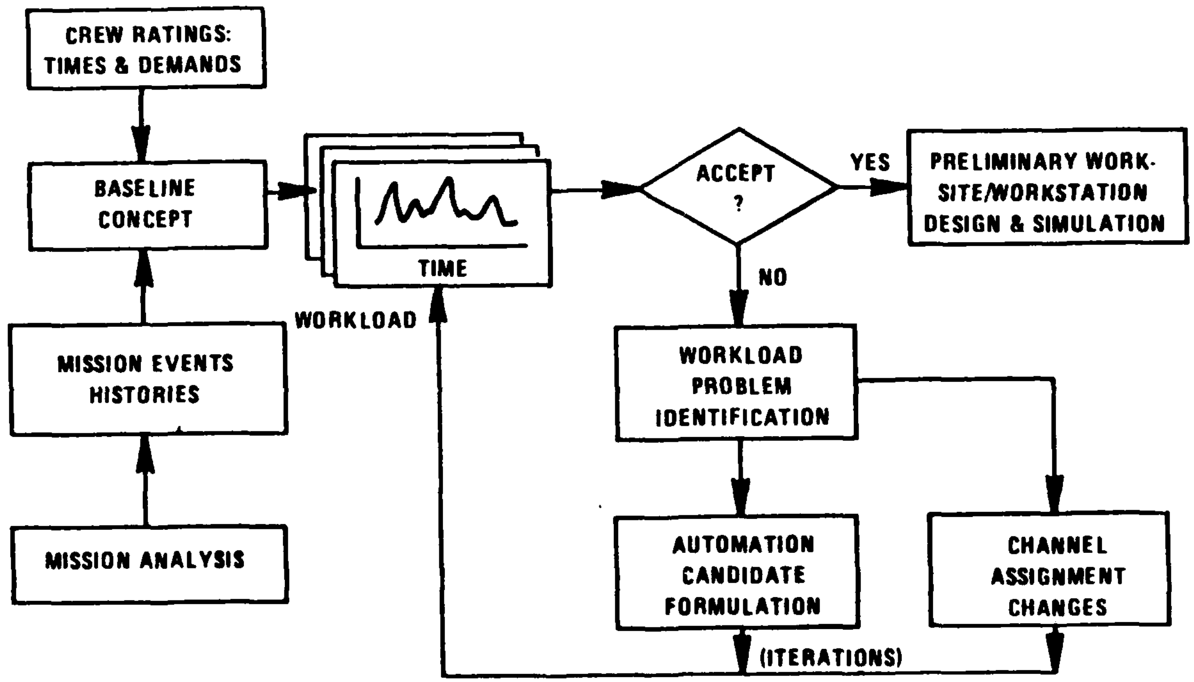
	VOC	VIC	VHC	AUD	MCO	MDC	SPE	COG
VIS OUTSIDE	0.5							
VIS INSIDE	0.9	0.5						
VIS HEAD	0.5	0.5	0.3					
AUDITORY	0.5	0.5	0.5	0.7				
MAN. CONT	0.1	0.1	0.1	0.1	0.9			
MAN. DISC	0.1	0.1	0.1	0.1	0.8	0.6		
SPEECH	0.1	0.1	0.1	0.4	0.3	0.1	1.0	
COGNITIVE*	0.7	0.7	0.7	0.9	0.1	0.3	0.8	1.0

*COGNITIVE ELEMENT UNRELATED TO CHANNEL IN CONFLICT

85V38

"WINDEX" WORKLOAD ANALYSIS PROCESS FLOW

Honeywell



85V38

Typical Interface/Activity Matrix, Battle Command SCAT

Honeywell

Activity Name		VOC	VIC	VHC	AUD	MCO	MDC	SPC	COG
FC	Bob Up	1	0	1	0	3	0	0	0
FC	Hover in ground effect	3	0	1	0	3	0	0	0
FC	Hover out of ground effect	1	0	2	0	3	0	0	0
FC	Nap of the earth	3	0	2	0	5	0	0	0
FC	Remask	3	0	2	0	5	0	0	0
FC	Select flight symbology	0	0	0	0	0	1	0	0
NV	Check navigation	0	2	0	0	0	0	0	0
NV	Update navigation	0	2	0	0	0	2	0	0
CO	Acknowledge receipt of data	0	0	0	1	0	0	0	0
CO	Data receive	0	0	0	2	0	0	0	0
CO	Data transmit	0	0	0	0	0	0	1	0
CO	Transmit battle position to team	0	0	0	0	0	0	2	0
ASE	Assess threats	0	0	0	0	0	0	0	3
ASE	Threat evasive maneuver	0	0	0	0	0	0	0	3
ASE	Threat response plan	0	0	0	0	0	0	0	4
TA	Acknowledge eye-sensor link	0	0	1	0	0	0	0	0
TA	Acknowledge sensor mode	0	1	0	0	0	0	0	0
TA	Acknowledge sensor ready	0	1	0	0	0	0	0	0
TA	Adjust sensor	0	0	1	0	1	0	0	0
TA	Link eye/sensor	0	0	0	0	0	1	0	0
TA	Select sensor	0	1	0	0	0	2	0	0
TA	Select sensor mode	0	1	0	0	0	2	0	0
TA	Slew sensor	1	0	1	0	3	0	0	0
TA	Target search	3	0	1	0	0	0	0	0
FRC	Acknowledge weapons elect	0	0	1	0	0	0	0	0
FRC	Select weapon	0	0	0	0	0	2	0	0
MM	Sensor management	0	0	0	0	0	0	0	2
MM	Stores management	0	0	0	0	0	0	0	3
MM	System monitor	0	0	0	0	0	0	0	2
MM	Tactical coordination	0	0	0	0	0	0	0	3
MM	Tactical decision	0	0	0	0	0	0	0	2

86-CRV-1596

CREW INTERFACE CHANNELS AND CONFLICTS

Honeywell

INTERFACE ABBREVIATION	FULL INTERFACE NAME
VOC	VISUAL OUT COCKPIT
VIC	VISUAL IN COCKPIT
VHC	VISUAL HEAD COUPLED
AUD	AUDITORY INPUT
MCO	MANUAL CONTINUOUS
MDC	MANUAL DISCRETE
SPC	SPEECH FROM PILOT
COG	COGNITIVE PROCESS

85V38

INSTANTANEOUS WORKLOAD DEMAND CALCULATION

Honeywell

$WD_{(i)}$ = (SUM OF CHANNEL/INTERFACE DEMANDS) +

(SUM OF CHANNEL/INTERFACE CONFLICT PENALTIES)

85V38

WORKLOAD ANALYSIS

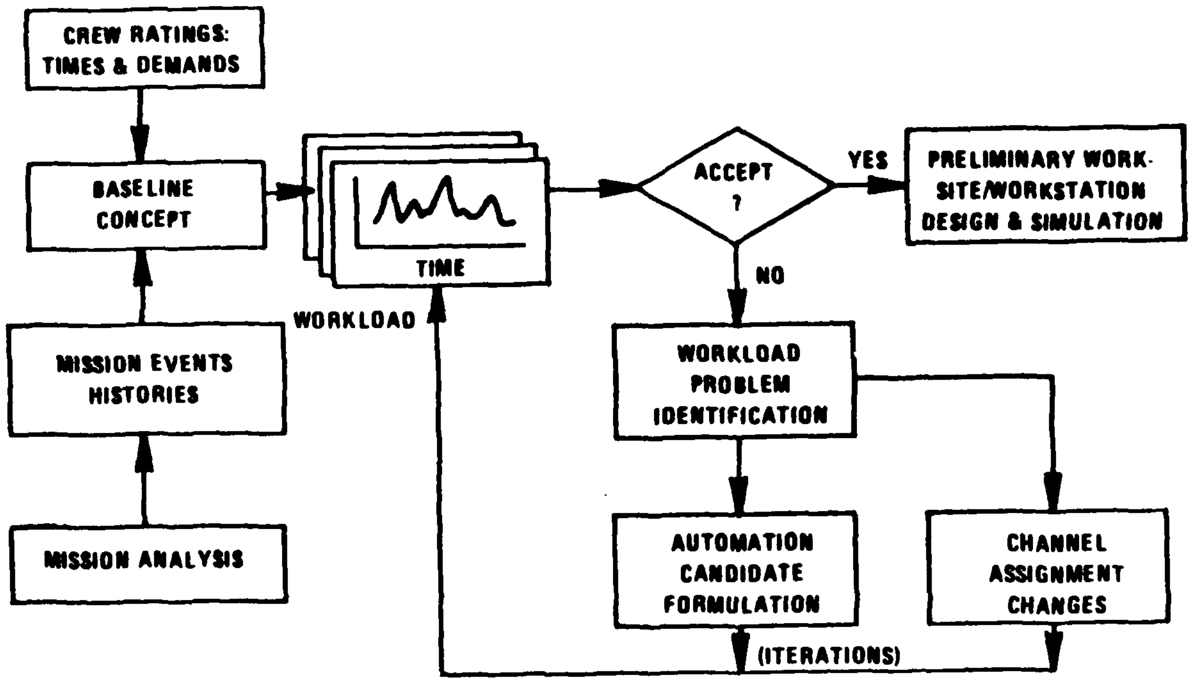
Honeywell

PURPOSE:

- ASSESS CREW WORKLOAD FOR SPECIFIED MISSION TASK
- IDENTIFY EXCESSIVE WORKLOAD "PEAKS"
- IDENTIFY FACTORS CONTRIBUTING TO EXCESSIVE WORKLOAD
- CONTROL ALLOCATION: AUTOMATE VERSUS MANUAL

85V38

"WINDEX" WORKLOAD ANALYSIS PROCESS FLOW Honeywell



85V38

WORKLOAD ANALYSIS CONTRACTS

Honeywell

PROGRAM NAME	AGENCY	TIME FRAME	RESULTS/SCOPE
PHYSIOLOGICAL MEASURES	(NASA/LRC, NAVY/NADC, ARMY/ECOM, AF/FDL)	1974-1978	HEART RATE, MUSCLE TENSION RESPIRATION RATE CORRELATES
ADVANCED DISPLAYS	AF/AMRL	1977-1978	PAPER ANALYSIS/"BEFORE & AFTER" DESIGNS
LHX	BOEING (ARMY)	1982	COMPUTERIZED TASK/WORKLOAD ESTIMATION
PAVE PILLAR	BOEING MAC (AFAC)	1983	COMPUTERIZED TASK/WORKLOAD ESTIMATION
ARTI/LHX	HUGHES HELICOPTERS	1984	AUTOMATION TASK IDENTIFICATION/ INTERFACE DESIGN

85V38

Four Workload Estimate Calculation Examples

Honeywell

Activities	Interfaces and Demands								Workload Components		
	VOC	VIC	VHC	AUD	MCC	MDC	SPE	COG	Linear	2nd-ord	Total
NOE Flight	3	0	2	0	5	0	0	0	10.0	4.0	14.0
NOE & Threat Assessment	3	0	2	0	5	0	0	3	13.0	12.5	25.5
NOE & Select Radio Channel	3	0	3	0	5	2	0	0	13.0	12.1	25.1
NOE & Select Radio Channel	3	0	3	0	5	0	2	0	13.0	8.6	21.6

86-CRV-1595

WORKLOAD ANALYSIS

Honeywell

WORKLOAD INDEX (WINDEX)

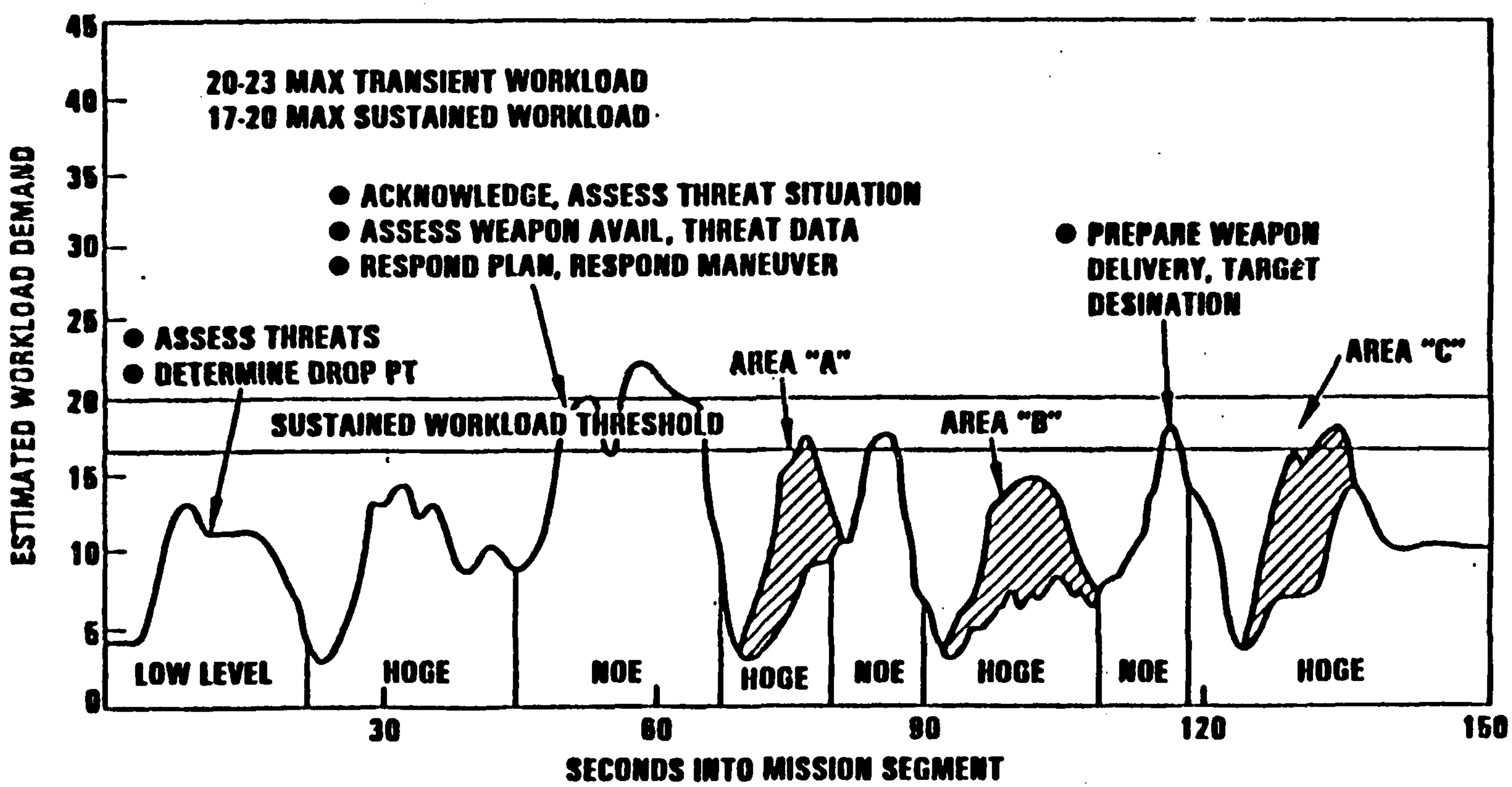
- MISSION/TASK ANALYSIS
- SUBJECT MATTER EXPERT INTERVIEWS
- ANALYSIS
- EXPERT REVIEW

85V38

	VOC	VIC	VHC	AUD	MCO	MDC	SPE	COG
VIS OUTSIDE	0.6							
VIS INSIDE	0.9	0.6						
VIS HEAD	0.5	0.5	0.3					
AUDITORY	0.6	0.6	0.6	0.7				
MAN. CONT	0.1	0.1	0.1	0.1	0.9			
MAN. DISC	0.1	0.1	0.1	0.1	0.8	0.6		
SPEECH	0.1	0.1	0.1	0.4	0.3	0.1	1.0	
COGNITIVE*	0.7	0.7	0.7	0.9	0.1	0.3	0.8	1.0

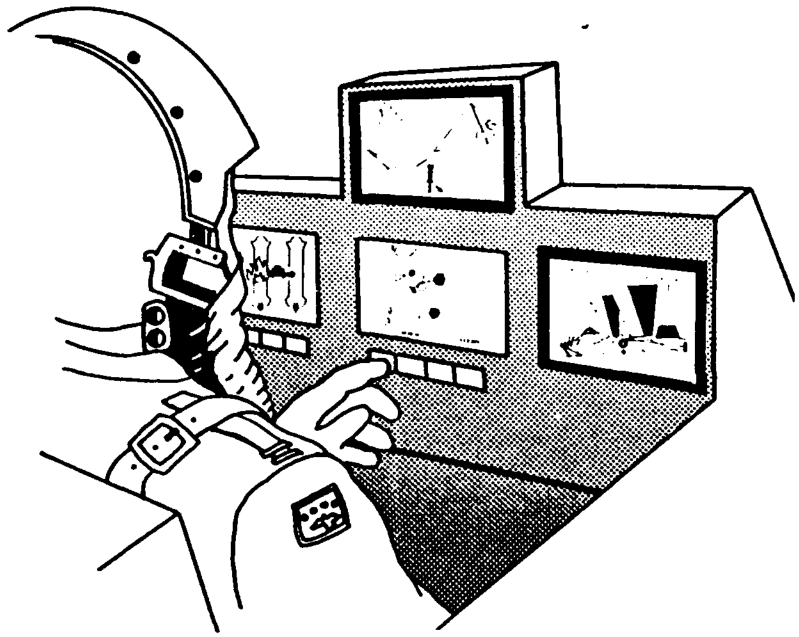
*COGNITIVE ELEMENT UNRELATED TO CHANNEL IN CONFLICT.

IMPACT OF VOICE COMMAND



Advanced Human Interface

Honeywell



Objective: Develop methods for optimizing flow of information input, processing, and output in DoD/NASA systems where the human plays critical role in system effectiveness.

- Pilot/Vehicle Interface
- ASW Work Station
- Workload Estimation Index
- Human Performance Laboratory
  - Part Task Simulation
  - Laser Protection Studies
- Cognitive Models

85-EGS-2705

ATTENDANCE LIST

Mr. Homer Adkins
CNET
Code N911
Naval Air Station
Pensacola, FL 32508
Av 922-4201

Mr. Robert Bateman
Boeing Military Airplane Co.
MS K 76-23
Wichita, KS 67210
(316) 526-2121

Mr. D. Beevis
DCIEM
P.O. Box 2000
Downsview, Ontario
CANADA M3M3B9
AV 827-4101

Dr. Mike Biferno
Mail Code 35-36
Douglas Aircraft Co.
3855 Lakewood Blvd.
Long Beach, CA 90840
(213) 593-6645

Mr. Harry Britten-Austin
Squadron Leader
AFWAL/FIGR
Wright-Patterson AFB, OH 45433
(513) 255-8259

Dr. Nancy M. Bucher
Research Manager
Crew Station Research and Development Office
TA-8
NASA-Ames Research Center
Moffett Field, CA 94035
(415) 694-4024

Mr. Larry Butterbaugh
AFWAL/FIGR
Wright-Patterson AFB, OH 45433
(513) 255-8260

Mr. Ron Cammarata
Director
USALABCOM
Human Engineering Laboratory
Attn: SLCHE-CC(Cammarata)
Aberdeen Proving Ground, MD 21005-5001
(301) 278-5950

Mr. Dave Carter
US Army ARL Commander
Box 577
SGRD-UAB-CS/Carter
Fort Rucker, AL 36362
AV 558-6862

Mr. Richard E. Christ
Chief
US Army Research Institute for the Behavioral
and Social Sciences-Ft. Bliss Field Unit
ATTN: Dr. Christ
P.O. Box 6057
Fort Bliss, TX 79906-0057
(915) 568-4491

Mr. Larry Clapp
Code 918
Naval Avionics Center
6000 E. 21st St.
Indianapolis, IN 46219-2189
(317) 351-4476

Eric Crawford
c/o Dr. Schiffler
ASD/ENECH
Wright-Patterson AFB, OH 45433

Mr. Pat Dowd
Crew Technology Division
USAFSAM/VN
Brooks AFB, TX 78235-5000
AV 240-3464

Dr. Douglas Eddy
Crew Technology Division
USAFSAM/VN
Brooks AFB, TX 78232-5000
AV 240-3464

Ms. DeAun Faulkner
Commander
Naval Air Test Center (SX726)
ATTN: DeAun Faulkner
Patuxent River, MD 20670-5304
AV 356-4157

Major Lucky Goebel
B-1B FOT & E ITDD
Dyess AFB, TX 79607-5270
AV 446-3514

Mr. Bruce Hamilton
78 Oxbow Lane
Northford, CT 06472
(203) 386-5834

Dr. Peter Hancock
University of Southern California
Department of Safety Science
ISSM, USC
Los Angeles, CA 90089-0021
(213) 743-2938

Dr. Randall L. Harris, Sr.
Langley Research Center
MS 152E
Hampton, VA 23665-5225
(804) 865-3917

Mr. K. C. Hendy
DCIEM
P.O. Box 2000
Downsview, Ontario
CANADA M3M3B9
Av 827-4101

Mr. Dick Horst
ARD Corp.
5457 Twin Knolls Rd.
Columbia, MD 21045
(301) 596-5845

Dr. Daniel Jones
USNRC
DHFT/HFIB(AR5209)
Washington, DC 20555
(301) 492-4846

Ms. Karen Jones
Coast Guard Institute
US Coast Guard
P.O. Sub Station 18
Oklahoma City, OK 73169
(405) 686-2416

Mr. Jerry Kennebrew
580 N. Park St. #10
Columbus, OH 43215
(614) 860-7193

Ms. June E. Lay
General Dynamics
P.O. Box 748
MZ 1766
Ft. Worth, TX 76101
(817) 763-2780

Mr. Kwan Lee
3128 Ceba Dr.
Louisiana State University
Baton Rouge, LA 70803
(504) 388-5369

Dr. Nita Lewis
Crew Technology Division
USAFSAM/VN
Brooks AFB, TX 78235-5000
AV 240-3464

Ms. Gretchen Lizza
c/o Dr. Schiffler
ASD/ENECH
Wright-Patterson AFB, OH 45433
(513) 237-0473

Mr. John Lockett
Director
USALABCOM
Human Engineering Laboratory
Attn: SLCHE-CC(Lockett)
Aberdeen Proving Ground, MD 21005-5001
(301) 278-5950

Mr. Daniel McCrobie
MZ 601-19
General Dynamics
P.O. Box 50-800
Ontario, CA 91761
(714) 945-7562

Mr. Lou Meyers
Battelle
505 King Avenue
Columbus, OH 43201
(614) 424-7166

Dr. James Miller
Crew Technology Division
USAFSAM/VN
Brooks AFB, TX 78235-5000
AV 240-3464

Ms. Andrea Morris
Boeing VERTOL
P.O. Box 16858
MS-P3032
Philadelphia, PA 19142
(215) 499-9635

Ms. Deborah Mummaw
6520 TESTG/ENAH
Edwards AFB, CA 93523-5000
(805) 277-3334

LT COL Mark Nataupsky
MS 152-E
NASA-Langley Research Center
Hampton, VA 23665-5225
(804) 865-3917

Ms. Laurie Quill
Beville Engineering, Inc.
270 Regency Ridge, Suite 214
Dayton, OH 45459
(513) 434-1093

Dr. Matilda Reeder
Army Research Institute
PERI-IR
Ft. Rucker, AL 36362-5000
AV 558-4204

Dr. Richard J. Schiffler
ASD/ENECH
Wright-Patterson AFB, OH 45433
(513) 255-5597

CAPT John K. Schmidt
Director
USALABCOM
Human Engineering Laboratory
Attn: SLCHE-CC(CAPT Schmidt)
Aberdeen Proving Ground, MD 21005-5001
(301) 278-5836

Dr. Larry Shadow
6545 Test Group/TIE
Hill AFB, UT 84056-5000
AV 458-6195

Ms. Kim Siler
414 E. Cedar, Apt. 3
Burbank, CA 91501
(818) 847-7591

Mr. Rom Simmons
US Army ARL Commander
Box 577
SGRD-UAB-CS/Carter
Fort Rucker, AL 36362
AV 558-6862

Ms. Mary Lou Smith
c/o Dr. Schiffler
ASD/ENECH
Wright-Patterson AFB, OH 45433
(513) 429-2252

Mr. Paul Stager
Department of Psychology
York University
4700 Keele Street
Toronto, Ontario
CANADA M3J1P3
(416) 736-2100

Mr. Lewis Stone
US Army ARL Commander
Box 577
SGRD-UAB-CS/Stone
Fort Rucker, AL 36362
AV 558-6853

Mr. Richard Tauson
Director
USA Human Engineering Laboratory
ATTN: AMXHE-FT (R. A. Tauson)
Aberdeen Proving Ground, MD 21005-5001
AV 298-5874

Mr. Marvin Thandsen
c/o Dr. Schiffler
ASD/ENECH
Wright-Patterson AFB, OH 45433
(513) 767-2691

Ms. Mary Thiessen
General Dynamics
P.O. Box 748
MZ 1766
Ft. Worth, TX 76101
(817) 763-2780

Ms. Maria L. Thomas
Walter Reed Army Institute of Research
Division of Neuropsychiatry
Department of Human Behavioral Biology
ATTN: SGRD-UWI-C
Washington, DC 20307-5100
(301) 427-5521

Ms. Pamela Tsang
MS 239-3
NASA-Ames Research Center
Moffett Field, CA 94035
(415) 694-6187

Mr. Michael Vidulich
MS 239-3
NASA-Ames Research Center
Moffett Field, CA 94035
(415) 694-6187

Mr. Bob Weneck
Eastman Kodak Co.
Ergonomics Group
Bldg 320, 2nd Floor
Kodak Park
Rochester, NY 14650
(716) 722-6973

347/348 Reverse Blank

Ms. Leslie Whitaker
Department of Psychology
University of Missouri St. Louis
St. Louis, MO 63121
(314) 553-5479

INITIAL DISTRIBUTION

USAF Academy (Attn: LT COL T. McCloy, LT COL W. Derrick)	2
NASA, Ames Research Center (Attn: S. Hart)	1
Wright-Patterson AFB (Attn: G. Reid, G. Wilson)	2
Honeywell, Inc., Minneapolis (Attn: R. North)	1
Northrop Corp., Aircraft Division, Hawthorne, CA (Attn: J. Armstrong)	1
U. of Dayton (Attn: T. Eggemeier)	1
Washington U., St. Louis (Attn: J. Stern)	1
JAYCOR, Fairborn, OH (Attn: T. Ramirez)	1
DTIC	12

END
12-87
DTIC